



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 184078

TO: Manjunath N Rao
Location: rem/2A01/2C70
Art Unit: 1652
Tuesday, ~~April 11, 2006~~
Case Serial Number: 10/784870

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

BOB
barbara.obryen@uspto.gov

Search Notes

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: April 7, 2006, 17:32:32 ; Search time 138.392 Seconds
(without alignments)
2028.756 Million cell updates/sec

Title: US-10-784-870-1

Perfect score: 3059
Sequence: 1 MRKKVFLSVLSAAAILSTV.....EYQAVNPVGPQXPSLATVN 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq.21:*
1: geneseqp1980a:*
2: geneseqp1990a:*
3: geneseqp2000a:*
4: geneseqp2001a:*
5: geneseqp2002a:*
6: geneseqp2003a:*
7: geneseqp2003b:*
8: geneseqp2004a:*
9: geneseqp2005a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3030	99.1	639	2	AA117089 Bacillus
2	3022	98.8	640	2	AA117090 Bacillus
3	3022	98.8	640	8	ADM40773 Alkaline
4	3022	98.8	640	8	ADSI14437 Bacillus
5	3022	98.8	641	8	ADT49613 Bacillus
6	3021	98.8	640	2	AA117091 Bacillus
7	3017	98.6	639	2	AA117087 Bacillus
8	3007	98.3	640	2	AA117088 An alkali
9	2723	89.0	641	2	AAW89547 Bacillus
10	2408	78.7	636	2	AAW89548 Bacillus
11	2155	70.4	434	5	AAW50080 Bacillus
12	2155	70.4	434	5	AAW50081 Bacillus
13	2155	70.4	434	7	ADY33778 Bacillus
14	2155	70.4	434	7	ADZ51758 Mutant Ba
15	2155	70.4	434	7	ADZ51757 Wild-type
16	2155	70.4	434	8	ADL25802 Bacillus
17	2155	70.4	434	8	ADM40779 Alkaline
18	2155	70.4	434	8	ADM40771 Mature al
19	2155	70.4	434	8	ADM40780 Alkaline
20	2155	70.4	434	8	ADM40787 Bacillus
21	2150	70.3	434	8	ADSI14427 Bacillus
22	2147	70.2	434	8	ADSI14441 Bacillus
23	2147	70.2	434	8	ADSI14438 Bacillus
24	2147	70.2	434	8	ADSI14442 Bacillus

25	2146	70.2	434	8	ADSI14439 Bacillus
26	2140	70.0	434	8	ADSI14440 Bacillus
27	2138	69.9	434	8	ADSI14443 Bacillus
28	2133	69.7	434	8	ADSI14444 Bacillus
29	2118	69.2	434	8	ADSI14445 Alkaline
30	2103	68.7	436	8	ADM40787 Bacillus
31	2082	68.1	434	5	AAW50085 Bacillus
32	2082	68.1	434	7	ADZ51762 Mutant Ba
33	2064.5	67.5	433	8	ADM40784 Alkaline
34	2064.5	67.5	433	8	ADSI2082 Bacillus
35	2061.5	67.4	433	8	ADSI2083 Bacillus
36	2060.5	67.4	433	5	AAW50086 Bacillus
37	2060.5	67.4	433	7	ADZ51763 Mutant Ba
38	2060.5	67.4	433	8	ADM40785 Alkaline
39	2060.5	67.4	433	8	ADSI2010 Bacillus
40	2060.5	67.4	433	8	ADSI2078 Bacillus
41	2060.5	67.4	433	8	ADSI2016 Bacillus
42	2060.5	67.4	433	8	ADSI2011 Bacillus
43	2060.5	67.4	433	8	ADSI2001 Bacillus
44	2059.5	67.3	433	8	ADSI2062 Bacillus
45	2058.5	67.3	433	8	ADSI2054 Bacillus

ALIGNMENTS

RESULT 1	AA117089	standard; protein; 639 AA.
ID	AA117089	standard; protein; 639 AA.
XX	XX	XX
AC	AA117089;	XX
XX	XX	XX
DT	20-MAR-2003 (revised)	XX
DT	21-JUL-1999 (first entry)	XX
XX	XX	XX
DE	Bacillus alkaline protease.	XX
KW	Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme; washing composition; oxidising agent.	XX
OS	Bacillus sp.	XX
PN	WO9918218-A1.	XX
XX	XX	XX
PD	15-APR-1999.	XX
XX	XX	XX
PF	07-OCT-1998; 98WO-JP004528.	XX
PR	07-OCT-1997; 97JP-00274570.	XX
XX	XX	XX
PA	(KAOS) KAO CORP.	XX
PI	Takaiwa M, Okuda M, Sasaki K, Kubota H, Hitomi J, Kageyama Y;	XX
PI	Shikata S, Nomura M,	XX
XX	WPI: 1999-287736/27.	XX
DR	N-PSDB; AAX37277.	XX
XX	Alkali protease from Bacillus used in washing powders.	XX
PT	Discloure; Page 53-58; 71pp; Japanese.	XX
PS	XX	XX
XX	XX	XX
CC	The invention relates to alkaline proteases produced by strains of Bacillus. The proteases ability to digest casein is not inhibited by oleic acid and they have a high stability to oxidising agents. The alkali protease of the invention has the following properties: (a) it is active over the pH range 4-13 and has at least 80% of its optimum activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d) its ability to digest casein is not inhibited by oleic acid; (e) it has molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be used as enzymes in washing compositions for use in automatic dishwashers and for washing clothes. The stability to oxidising agents allows the	ADSI14443 Bacillus ADSI14440 Bacillus ADSI14443 Bacillus ADSI14444 Bacillus ADSI14445 Alkaline ADM40787 Bacillus AAW50085 Bacillus ADM40784 Alkaline ADSI2082 Bacillus ADSI2083 Bacillus AAW50086 Bacillus ADZ51763 Mutant Ba ADM40785 Alkaline ADSI2010 Bacillus ADSI2078 Bacillus ADSI2016 Bacillus ADSI2011 Bacillus ADSI2001 Bacillus ADSI2062 Bacillus ADSI2054 Bacillus

CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)
XX

Sequence 639 AA;

Query Match 99.1%; Score 3030; DB 2; Length 639;
Best Local Similarity 93.4%; Pred. No. 3.4e-237;
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
QY 1 MRKKCVFLSVLSAAAILSTVALKNPSAGAXRFDLDFKGIQTTTDXKGFSGKXQTGAAPL 60
DB 1 MRKKCVFLSVLSAAAILSTVALKNPSAGAXRFDLDFKGIQTTTDXKGFSGKXQTGAAPL 60
QY 61 IVESENVKLLKGLKLETVPAKKLHIQFNGPIIEETKOKLEKGAIIIDYIPYAI 120
DB 61 IVESENVKLLKGLKLETVPAKKLHIQFNGPIIEETKOKLEKGAIIIDYIPYAI 120
QY 121 VVEYGDVSKXXXIIEHVESVEPYLPYXIIDPOLFTKGASXLVKAXALDTKXNKKEVQLRG 180
DB 121 VVEYGDVSKXXXIIEHVESVEPYLPYXIIDPOLFTKGASXLVKAXALDTKXNKKEVQLRG 180
QY 181 IEXIAQXXXNDVYITTAPEYKVMNDVARGIVKADVAQSYGLYGGQIIVAVDTGLDT 240
DB 181 IEXIAQXXXNDVYITTAPEYKVMNDVARGIVKADVAQSYGLYGGQIIVAVDTGLDT 240
QY 241 GRNDSMHEAFRGKITLALYALGRTNANDTNGHGTVAAGVYGGNGTNGMAPOANLVFQ 300
DB 241 GRNDSMHEAFRGKITLALYALGRTNANDTNGHGTVAAGVYGGNGTNGMAPOANLVFQ 300
QY 301 SIMDSXGIGGLPSNLQTLFSGQXASAGARIHTNSWGAANVGAVTTDSRVDDYVRKNDT 360
DB 301 SIMDSXGIGGLPSNLQTLFSGQXASAGARIHTNSWGAANVGAVTTDSRVDDYVRKNDT 360
QY 361 IIPAGNEKPNNGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKQGR 420
DB 361 IIPAGNEKPNNGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKQGR 420
QY 421 IKPDVMAFGTYILSARSSLAAPDSFMANHDSKAYVMGTSMAPIVAGVNAQLREHPVKN 480
DB 421 IKPDVMAFGTYILSARSSLAAPDSFMANHDSKAYVMGTSMAPIVAGVNAQLREHPVKN 480
QY 481 RGITRPSILKALIAAGADVGLGYPNGQGWGRVTLDRKSLNVAIVNSSLSTSQKATY 540
DB 481 RGITRPSILKALIAAGADVGLGYPNGQGWGRVTLDRKSLNVAIVNSSLSTSQKATY 540
QY 541 XETATAGRPDKTISLWSAPASTTASVTLVNDLIVTPANGTYVGGDPKXPKXXANDG 600
DB 541 XETATAGRPDKTISLWSAPASTTASVTLVNDLIVTPANGTYVGGDPKXPKXXANDG 600
QY 601 RNNVENVFINKPQSGTYTIEVQAVNVVPGQXFSIAIYN 639
DB 601 RNNVENVFINKPQSGTYTIEVQAVNVVPGQXFSIAIYN 639
```

RESULT 2
AA17090 standard; protein; 640 AA.

```
XX AA17090;
XX
XX 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
XX DE Bacillus alkaline protease.
XX
XX KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX washing composition; oxidizing agent.
XX
XX Bacillus sp.
XX
XX MO9918218-A1.
```

PD 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
PF
XX
XX 07-OCT-1997; 97JP-00274570.
PR
XX
XX (KAOS) KAO CORP.
PI
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX
XX WPI: 1999-287736/27.
DR N-PSDB; JAX37278.

Alkali protease from Bacillus used in washing powders.
Dieckmann; Page 58-63; 71pp; Japanese.

The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)
XX

Sequence 640 AA;

Query Match 98.8%; Score 3022; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 1.5e-236;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```
QY 2 RKKCVFLSVLSAAAILSTVALKNPSAGAXRFDLDFKGIQTTTDXKGFSGKXQTGAAPL 61
DB 3 RKKCVFLSVLSAAAILSTVALKNPSAGAXRFDLDFKGIQTTTDXKGFSGKXQTGAAPL 62
QY 62 VESBNVLLKGLKLETVPAKKLHIQFNGPIIEETKOKLEKGAIIIDYIPYAI 121
DB 62 VESBNVLLKGLKLETVPAKKLHIQFNGPIIEETKOKLEKGAIIIDYIPYAI 122
QY 122 EYEGDVSKXXXIIEHVESVEPYLPYXIIDPOLFTKGASXLVKAXALDTKXNKKEVQLRG 181
DB 122 EYEGDVSKXXXIIEHVESVEPYLPYXIIDPOLFTKGASXLVKAXALDTKXNKKEVQLRG 182
QY 181 EYEGDVSKXXXIIEHVESVEPYLPYXIIDPOLFTKGASXLVKAXALDTKXNKKEVQLRG 182
DB 181 EYEGDVSKXXXIIEHVESVEPYLPYXIIDPOLFTKGASXLVKAXALDTKXNKKEVQLRG 182
QY 241 EXIAQXXXNDVYITTAPEYKVMNDVARGIVKADVAQSYGLYGGQIIVAVDTGLDTG 241
DB 241 EXIAQXXXNDVYITTAPEYKVMNDVARGIVKADVAQSYGLYGGQIIVAVDTGLDTG 242
QY 242 RNDSSMHEAFRGKITLALYALGRTNANDTNGHGTVAAGVYGGNGTNGMAPOANLVFQ 301
DB 242 RNDSSMHEAFRGKITLALYALGRTNANDTNGHGTVAAGVYGGNGTNGMAPOANLVFQ 302
QY 302 IMDSXGIGGLPSNLQTLFSGQXASAGARIHTNSWGAANVGAVTTDSRVDDYVRKNDT 361
DB 302 IMDSXGIGGLPSNLQTLFSGQXASAGARIHTNSWGAANVGAVTTDSRVDDYVRKNDT 362
QY 362 LFAAGNEKPNNGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKQGR 421
DB 362 LFAAGNEKPNNGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKQGR 422
QY 421 KPDVMAFGTYILSARSSLAAPDSFMANHDSKAYVMGTSMAPIVAGVNAQLREHPVKN 481
DB 421 KPDVMAFGTYILSARSSLAAPDSFMANHDSKAYVMGTSMAPIVAGVNAQLREHPVKN 482
QY 482 GITRPSILKALIAAGADVGLGYPNGQGWGRVTLDRKSLNVAIVNSSLSTSQKATY 541
DB 482 GITRPSILKALIAAGADVGLGYPNGQGWGRVTLDRKSLNVAIVNSSLSTSQKATY 541
```

```

Db      483 GTPPKSLKALIALAGADIGLGYPNGNGMGWRTLDKSLNVAYNNESSLSTSQKATYS 542
QY      542 PTATGKPLKISLWSDAPASTTASTVTLVNDLVLVTAPNGTXYGNDPXXPKXXWDR 601
Db      543 PTATGKPLKISLWSDAPASTTASTVTLVNDLVLVTAPNGTXYGNDPXXPKXXWDR 602
QY      602 NNVENFINKPQSGTYTIEVQAYNVPVGPQXPSLAIVN 639
Db      603 NNVENFINAPQSGTYTIEVQAYNVPVGPQXPSLAIVN 640

RESULT 3
ADM40773
ID      ADM40773 standard; protein; 640 AA.
XX
AC      ADM40773;
XX
DT      01-JUL-2004 (first entry)
XX
DE      Alkaline protease from Bacillus sp. KSM-KP43.
XX
KM      alkaline protease; laundry detergent; bleaching agent; detergent;
XX      denture-cleaning agent; enzyme.
XX
OS      Bacillus sp.; KSM-KP43.
XX
PN      US2004072321-A1.
XX
PD      15-APR-2004.
XX
PF      09-JUN-2003; 2003US-00456479.
XX
PR      26-JUN-2002; 2002JP-00186387.
XX      18-OCT-2002; 2002JP-00304232.
XX      (KAOS ) KAO CORP.
XX
PI      Sato T, Okuda M, Takimura Y, Sumitomo N, Nomura M, Kobayashi T;
XX
DR      WPI, 2004-328572/30.
XX
DR      N-PSDB; ADM40772.
XX
PT      New alkaline protease having a mutant prepro sequence where amino acid
PT      residues at positions 52, 75 and 142 are substituted with another amino
PT      acid residue, useful as enzyme component of laundry detergents, or
PT      bleaching agents.
XX
PS      Discloure; SEQ ID NO 4; 29pp; English.
XX
XX
CC      The invention relates to an alkaline protease having a prepro sequence.
CC      The prepro sequence is a mutant sequence of SEQ ID NO: 1, or an amino
CC      acid sequence having 80% homology or higher to SEQ ID NO: 1, where amino
CC      acid residues at: (a) position 52 is substituted by aspartic acid or
CC      arginine; (b) position 75 is substituted by alanine or arginine; and (c)
CC      position 142 is substituted by lysine. The alkaline protease is useful as
CC      an enzyme which can be incorporated into laundry detergents, bleaching
CC      agents, detergents for cleaning hard surfaces or drainpipes, denture-
CC      cleaning agents, and detergents for sterilizing medical apparatus. The
CC      present sequence represents alkaline protease from Bacillus sp. KSM-KP43.
XX
SQ      Sequence 640 AA;

Query Match      98.8%; Score 3022; DB 8; Length 640;
Best Local Similarity 93.3%; Pred. No. 1.5e-236;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```

```

QY      122 EYEGDVKXXXXIIEHVESVEPYLPKXXIDPOLPTKGASLYVRKXALDTKQXKKEVQLRGI 181
Db      123 EYEGDVKASTSTIEHVESVEPYLPYRIDPOLFTKGASLVAVALDTQKKEVQLRGI 182
QY      182 EXIAQKXKXNDVXYITAKPEYKVMNDVARGIKVADVAQSSYGLYGGGQIIVAADTGLDYG 241
Db      183 EQIAQPAISNDVLYITAKPEYKVMNDVARGIKVADVAQSSYGLYGGGQIIVAADTGLDYG 242
QY      242 RNDSSMHEAFRGKITALVALGRTNNANDTNGHGTIVAGSVLGNKGTNKGAPQANLVFQS 301
Db      243 RNDSSMHEAFRGKITALVALGRTNNANDTNGHGTIVAGSVLGNKGTNKGAPQANLVFQS 302
QY      302 IMDSXGGLGGLPSNIQTLPSQAXSAGARHTMSKGAANVGAATTTDSRNDVYRKNDMTI 361
Db      303 IMDSGGGLGGLPSNIQTLPSQAYAGARHTMSKGAANVGAATTTDSRNDVYRKNDMTI 362
QY      362 LPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQPSRGPTRDGR 421
Db      363 LPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQPSRGPTRDGR 422
QY      422 KPDVNAFGTXILSARSSLPDSSFWANHDSKYAYMGTSMAATPIVAGNVAQLRBEHFVQR 481
Db      423 KPDVNAFGTXILSARSSLPDSSFWANHDSKYAYMGTSMAATPIVAGNVAQLRBEHFVQR 482
QY      482 GTPPKSLKALIALAGADIGLGYPNGNGMGWRTLDKSLNVAYNNESSLSTSQKATYX 541
Db      483 GTPPKSLKALIALAGADIGLGYPNGNGMGWRTLDKSLNVAYNNESSLSTSQKATYS 542
QY      542 PTATGKPLKISLWSDAPASTTASTVTLVNDLVLVTAPNGTXYGNDPXXPKXXWDR 601
Db      543 PTATGKPLKISLWSDAPASTTASTVTLVNDLVLVTAPNGTXYGNDPXXPKXXWDR 602
QY      602 NNVENFINKPQSGTYTIEVQAYNVPVGPQXPSLAIVN 639
Db      603 NNVENFINAPQSGTYTIEVQAYNVPVGPQXPSLAIVN 640

RESULT 4
ADS14437
ID      ADS14437 standard; protein; 640 AA.
XX
AC      ADS14437;
XX
DT      30-DEC-2004 (first entry)
XX
DE      Bacillus alkaline protease KP43.
XX
KM      protease; enzyme; alkaline protease; laundry detergent; KP43.
XX
OS      Bacillus sp.; KSM-KP43.
XX
FH      Key Location/Qualifiers
PT      Peptide 1..206
PT      Protein 207..640
XX
PN      EP1466962-A1.
XX
PD      13-OCT-2004.
XX
PP      08-APR-2004; 2004EP-00008604.
XX
PR      10-APR-2003; 2003JP-00106709.
XX
PA      (KAOS ) KAO CORP.
XX
PI      Okuda M, Sato T, Takimura Y, Sumitomo N, Kobayashi T;
XX
DR      WPI, 2004-713113/70.
XX
DR      N-PSDB; ADS14428.
XX
PT      New engineered alkaline protease, useful particularly in laundry
PT      detergents, comprising specified amino acids at particular positions.

```

XX Disclosure; SEQ ID NO 1; 31pp; English.
PS
XX
CC The invention relates to a novel alkaline protease. The new alkaline
CC protease comprises an amino acid sequence in which one or more amino acid
CC residues selected from those located at 7 specific positions within
CC ADS14427, or at positions corresponding to these positions are: position
CC 15 (histidine), position 16 (threonine or glutamine), position 166
CC (glycine), position 167 (valine), position 167 (serine), position 346
CC (arginine), and position 405 (aspartic acid). The alkaline protease is
CC useful in industry particularly in laundry detergents, but also e.g. in
CC fibre modifying agents, leather processing agents, cosmetic compositions,
CC bath additives, food-modifying agents, and pharmaceuticals. The present
CC sequence represents the wild-type Bacillus sp. KSM-KP43 alkaline
CC protease, KP43. The sequence is shown in the sequence listing as part of
CC SEQ ID NO:1
XX
SQ Sequence 640 AA;
Query Match 98.8%; Score 3022; DB 8; Length 640;
Best Local Similarity 93.3%; Pred. No. 1.5e-236;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
QY 2 RKKKVFSLVSAAILSTVALNPSAGKARXFDLDFKGIQTTTDXKXGSKOQTGAAPL 61
DB 3 KKKKVFSLVSAAILSTVALNPSAGARNDLDFKGIQTTTDAKGFSGKOGTGAAPL 62
QY 62 VSEENVKLPKGLQKLETVANNKLIHQPNGBILLETQXLETKAKLIDYIPVAYIV 121
DB 63 VSEENVKLPKGLQKLETVANNKLIHQPNGBILLETQXLETKAKLIDYIPVAYIV 122
QY 122 EYEGDVYKSAITIEHVESVEPYLPXYXIDPOLFTKASXLVKAALDTQXNKEVQLRGI 181
DB 123 EYEGDVYKSAITIEHVESVEPYLPYRIPDLFTKASLVAVALDTQXNKEVQLRGI 182
QY 182 EXIAQXXXNDVYITTAKEPYKMDVARGIVADVAQSSYGLYGGQIVAAVDTGLDTG 241
DB 183 EOIAPFALSNDVLYITAKEPYKMDVARGIVADVAQSSYGLYGGQIVAAVDTGLDTG 242
QY 242 RNDSSMHEAFRGKITALVYALGRTNNANDNGHGTVAAGSVLGNKTNKGMAPQANLVFQS 301
DB 243 RNDSSMHEAFRGKITALVYALGRTNNANDNGHGTVAAGSVLGNKTNKGMAPQANLVFQS 302
QY 302 IMDSXGGLGSLPSNLQTLFQXASAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 361
DB 303 IMDSXGGLGSLPSNLQTLFQXASAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 362
QY 362 LPAAGNHEPNGGTISAPGTAKNAITVGTATENLRPSFGSYADNINHVAQFSSRGPTKGR 421
DB 363 LPAAGNHEPNGGTISAPGTAKNAITVGTATENLRPSFGSYADNINHVAQFSSRGPTKGR 422
QY 422 KPDVWAPGTILISARSLAPDSSFWANHSKYAVNGGTSMATPIYAGVAVQALREHFVNR 481
DB 423 KPDVWAPGTILISARSLAPDSSFWANHSKYAVNGGTSMATPIYAGVAVQALREHFVNR 482
QY 482 GTPKPSLLKALIGAADIGLGYNGNGQWGRVTLDSLNVAYVNESSSLSTOKATYS 541
DB 483 GTPKPSLLKALIGAADIGLGYNGNGQWGRVTLDSLNVAYVNESSSLSTOKATYS 542
QY 542 FTATAGKPLKISLVWSDAPASTTASVTLVNDLVTTPNGTXYVGNDFXXPKXNMNDR 601
DB 543 FTATAGKPLKISLVWSDAPASTTASVTLVNDLVTTPNGTXYVGNDFTPYNDNMGR 602
QY 602 NNVENVFINKPOSQTYTIEVOAYVNPVGPQKFSLAIVN 639
DB 603 NNVENVFINKPOSQTYTIEVOAYVNPVGPQKFSLAIVN 640
RESULT 5
ADT49613
XX ADT49613 standard; protein; 641 AA.
XX ADT49613;
AC

XX 30-DEC-2004 (first entry)
DT
XX Bacillus alkaline protease.
XX
DB Alkaline protease; detergent; fiber modification; leather processing;
KW cosmetic; bath additives; food-modification; pharmaceutical; enzyme.
XX Bacillus sp. KSM-KP43.
OS
XX
XX
XX Key Location/Qualifiers
XX Peptide 1..206
XX Protein /note= "signal peptide"
XX /note= "specifically claimed mature protein (SEQ ID 1)"
XX
XX EPI466970-AI.
XX
XX
XX 13-OCT-2004.
XX
XX 08-APR-2004; 2004EP-00008605.
XX
XX 10-APR-2003; 2003JP-00106708.
XX
XX (KAOS) KAO CORP.
XX
XX Sato T, Okuda M, Koyama S, Izawa Y, Kobayashi T,
XX WPI; 2004-711317/70.
XX DR N-PSDB; ADT49605.
XX
XX New engineered alkaline protease with improved activity and thermal
XX stability, useful particularly in detergents such as laundry detergents.
XX
XX Disclosure; Page 19-25; 35pp; English.
XX
XX The invention relates to an alkaline protease and its encoding gene. The
XX alkaline protease can be expressed by standard recombinant methodology.
XX The alkaline protease is useful in the industry particularly in
XX detergents such as laundry detergents, but also in fiber modifying
XX agents, leather processing agents, cosmetic compositions, bath additives,
XX food-modifying agents, and pharmaceuticals. The enzyme has good activity
XX and thermal stability. The present sequence represents an alkaline
XX protease from Bacillus sp. KSM-KP43.
XX
XX Sequence 641 AA;
SQ
Query Match 98.8%; Score 3022; DB 8; Length 641;
Best Local Similarity 93.3%; Pred. No. 1.5e-236;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
QY 2 RKKKVFSLVSAAILSTVALNPSAGKARXFDLDFKGIQTTTDXKXGSKOQTGAAPL 61
DB 3 KKKKVFSLVSAAILSTVALNPSAGARNDLDFKGIQTTTDAKGFSGKOGTGAAPL 62
QY 62 VSEENVKLPKGLQKLETVANNKLIHQPNGBILLETQXLETKAKLIDYIPVAYIV 121
DB 63 VSEENVKLPKGLQKLETVANNKLIHQPNGBILLETQXLETKAKLIDYIPVAYIV 122
QY 122 EYEGDVYKSAITIEHVESVEPYLPXYXIDPOLFTKASXLVKAALDTQXNKEVQLRGI 181
DB 123 EYEGDVYKSAITIEHVESVEPYLPYRIPDLFTKASLVAVALDTQXNKEVQLRGI 182
QY 182 EXIAQXXXNDVYITTAKEPYKMDVARGIVADVAQSSYGLYGGQIVAAVDTGLDTG 241
DB 183 EOIAPFALSNDVLYITAKEPYKMDVARGIVADVAQSSYGLYGGQIVAAVDTGLDTG 242
QY 242 RNDSSMHEAFRGKITALVYALGRTNNANDNGHGTVAAGSVLGNKTNKGMAPQANLVFQS 301
DB 243 RNDSSMHEAFRGKITALVYALGRTNNANDNGHGTVAAGSVLGNKTNKGMAPQANLVFQS 302
QY 302 IMDSXGGLGSLPSNLQTLFQXASAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 361

```
Db      303  IMDSGGGLGGLPSNLQTLFSGQAYSAGARIHTNSGMAAVNGAYTTDSRNVDDYVRKNDMTI 362
Qy      362  LPAAGNEKPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDGR1 421
      363  LPAAGNEKPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDGR1 422
Qy      422  KPDVMAPGTXILISARSLAPDSSPFMANHDSKVAIYMGSTMAATPIVAGNVQALREHVVKNR 481
      423  KPDVMAPGTXILISARSLAPDSSPFMANHDSKVAIYMGSTMAATPIVAGNVQALREHVVKNR 482
Qy      482  GITPKPSLLKALILAGADXLGYPNGNGQMGWRVTLDKSLNVAIYVNESSXLSTSQKATYX 541
      483  GITPKPSLLKALILAGADXLGYPNGNGQMGWRVTLDKSLNVAIYVNESSXLSTSQKATYX 542
Qy      542  FTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGIXYVGNDFPKXPXKNWDCR 601
      543  FTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGIXYVGNDFPKXPXKNWDCR 602
Qy      602  NNVENVFINKPQSGTITIEVOAYNVPGQPFSLAIYN 639
      603  NNVENVFINKPQSGTITIEVOAYNVPGQPFSLAIYN 640
Db
```

```
RESULT 6
AA17091
ID  AA17091 standard; protein; 640 AA.
```

```
XX      AA17091;
AC      20-MAR-2003 (revised)
DT      21-JUL-1999 (first entry)
XX      Bacillus alkaline protease.
XX      Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX      washing composition; oxidising agent.
XX      Bacillus sp.
XX      WO9118218-A1.
XX      15-APR-1999.
XX      07-OCT-1998; 98WO-JP004528.
XX      07-OCT-1997; 97JP-00274570.
XX      (KAOS ) KAO CORP.
XX      Takaiwa M, Okuda M, Saeki K, Kubota H, Hltomi J, Kageyama Y;
XX      Shikata S, Nomura M;
XX      WPI: 1999-287736/27.
XX      N-PSDB; AAK37279.
PT      Alkali protease from Bacillus used in washing powders.
XX      Disclosure; Page 63-68; 71pp; Japanese.
XX      The invention relates to alkaline proteases produced by strains of
XX      Bacillus. The proteases ability to digest casein is not inhibited by
XX      oleic acid and they have a high stability to oxidising agents. The
XX      alkaline protease of the invention has the following properties: (a) it
XX      is active over the pH range 4-13 and has at least 80% of its optimum
XX      activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
XX      stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
XX      its ability to digest casein is not inhibited by oleic acid; (e) it has
XX      molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
XX      used as enzymes in washing compositions for use in automatic dishwashers
XX      and for washing clothes. The stability to oxidising agents allows the
XX      enzyme to be an effective component of washing compositions including
XX      bleaches. The present sequence represents an alkaline protease. (Updated
XX      on 20-MAR-2003 to correct DR field.)
```

```
XX      SQ      Sequence 640 AA;
```

```
Query Match          98.8%; Score 3021; DB 2; Length 640;
Best local Similarity 93.3%; Pred. No. 1, 8e-236;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
```

```
Qy      2  RKKRVFLSTLSAAIISTVALXNPSAGKARXPDLPDKGIQTTTDXKGFSGKQGTGAAPL 61
      3  KKKRVFLSTLSAAIISTVALXNPSAGKARXPDLPDKGIQTTTDAAGFSGKQGTGAAPL 62
Db      62  VSESENVKJLKKGKKLLETVPANNKLIHQPNGBIEETQXLEXTGAKTLDYIPDRAIYI 121
      63  VSESENVKJLKKGKKLLETVPANNKLIHQPNGBIEETQXLEXTGAKTLDYIPDRAIYI 122
Qy      122  EYEGDVSKXXXIIEHVESVEPYLPYXXIDPOLFTKASGLVAKALDTRQXNKEVOLRG1 181
      123  EYEGDVSKATSTIEHVESVEPYLPYRIDPOLFTKASGLVAKALDTRQXNKEVOLRG1 182
Db      182  EXIAQKKXNDVXYITAKEPEYKMDVANGIYKADVAQSSYGLYGGGQIVANADTGLDTG 241
      183  EQIAQPAISNDVLYITAKEPEYKMDVANGIYKADVAQSSYGLYGGGQIVANADTGLDTG 242
Qy      242  RNDSSMHEAFRGKITLVALGRTNANDNGHGHVAGSVLNGXTNKGMAPOANLVPQS 301
      243  RNDSSMHEAFRGKITLVALGRTNANDNGHGHVAGSVLNGXTNKGMAPOANLVPQS 302
Qy      302  IMDSXGGLGGLPSNLQTLFSGQAYSAGARIHTNSGMAAVNGAYTTDSRNVDDYVRKNDMTI 361
      303  IMDSGGGLGGLPSNLQTLFSGQAYSAGARIHTNSGMAAVNGAYTTDSRNVDDYVRKNDMTI 362
Qy      362  LPAAGNEKPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDGR1 421
      363  LPAAGNEKPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDGR1 422
Qy      422  KPDVMAPGTXILISARSLAPDSSPFMANHDSKVAIYMGSTMAATPIVAGNVQALREHVVKNR 481
      423  KPDVMAPGTXILISARSLAPDSSPFMANHDSKVAIYMGSTMAATPIVAGNVQALREHVVKNR 482
Qy      482  GITPKPSLLKALILAGADXLGYPNGNGQMGWRVTLDKSLNVAIYVNESSXLSTSQKATYX 541
      483  GITPKPSLLKALILAGADXLGYPNGNGQMGWRVTLDKSLNVAIYVNESSXLSTSQKATYX 542
Qy      542  FTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGIXYVGNDFPKXPXKNWDCR 601
      543  FTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGIXYVGNDFPKXPXKNWDCR 602
Qy      602  NNVENVFINKPQSGTITIEVOAYNVPGQPFSLAIYN 639
      603  NNVENVFINKPQSGTITIEVOAYNVPGQPFSLAIYN 640
Db
```

```
RESULT 7
AA17087
ID  AA17087 standard; protein; 639 AA.
```

```
XX      AA17087;
AC      20-MAR-2003 (revised)
DT      21-JUL-1999 (first entry)
XX      An alkaline protease sequence from Bacillus species.
XX      Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX      washing composition; oxidising agent.
XX      Bacillus sp.
XX      Key      Location/Qualifiers
XX      MISC-difference 1..639
XX      FT      /note="all residues indicated as Xaa are arbitrary amino
XX      acids"
```

```
PN MO9918218-A1.
XX
XX 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX (KAOS ) KAO CORP.
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
PI Shikata S, Nomura M;
DR WPI; 1999-287736/27.
XX N-PSDB; AAX37277.
XX
XX Alkali protease from Bacillus used in washing powders.
XX
XX Claim 3; Page 47-50; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease of the
CC invention. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 639 AA;

Query Match      98.6%; Score 3017; DB 2; Length 639;
Best Local Similarity 100.0%; Pred. No. 3.8e-236;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRKKVFLSVLSAAAILSTVALXNPSAGAXXFDLDFKGIQTTTDXGFSKQXQTGAAP 60
DB 1 MRKKVFLSVLSAAAILSTVALXNPSAGAXXFDLDFKGIQTTTDXGFSKQXQTGAAP 60
QY 61 LVSEENVLKLXGKGLKXKLETPVANNKLTHTXQFNGPILEETKQKLETTGAKIIDYIPDAYI 120
DB 61 LVSEENVLKLXGKGLKXKLETPVANNKLTHTXQFNGPILEETKQKLETTGAKIIDYIPDAYI 120
QY 121 VEYEGDVASXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVKAXALDTKOKNKEVOLRG 180
DB 121 VEYEGDVASXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVKAXALDTKOKNKEVOLRG 180
QY 181 IEXIAQXXXSNDVXTTAKPEYKVNNDVARGIVKADVAQSSYGLXGQGIYAVADTGLDT 240
DB 181 IEXIAQXXXSNDVXTTAKPEYKVNNDVARGIVKADVAQSSYGLXGQGIYAVADTGLDT 240
QY 241 GRNDSMHEAFRGKLTALYALGRTNNADTNGHGHVAGSVLGNKXTKTKGAPOANLVFQ 300
DB 241 GRNDSMHEAFRGKLTALYALGRTNNADTNGHGHVAGSVLGNKXTKTKGAPOANLVFQ 300
QY 301 SIMDSXGGLGGLPSNLQTLFSGQASAGARIHTNSWGAIVNGAYTTTDSRNDVDYRXNDMT 360
DB 301 SIMDSXGGLGGLPSNLQTLFSGQASAGARIHTNSWGAIVNGAYTTTDSRNDVDYRXNDMT 360
QY 361 ILFPAAGNEXPNNGTISAFGTAKNAITVGATENLRPSFGSYADNINMFAQFSSRPGTKGR 420
DB 361 ILFPAAGNEXPNNGTISAFGTAKNAITVGATENLRPSFGSYADNINMFAQFSSRPGTKGR 420
QY 421 IKPDVMAFGTITLSRSSLAPDSSFWAMHDSKYAMGTSMAATPIVAGNVAQLBHFVKN 480
DB 421 IKPDVMAFGTITLSRSSLAPDSSFWAMHDSKYAMGTSMAATPIVAGNVAQLBHFVKN 480
```

```
QY 481 RGITPKPSLLKRALIAGAADGICYPNCGWGRVTLDKSLNVAVNNESSXLSTSQKATY 540
DB 481 RGITPKPSLLKRALIAGAADGICYPNCGWGRVTLDKSLNVAVNNESSXLSTSQKATY 540
QY 541 XFTATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXPPXXMDG 600
DB 541 XFTATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXPPXXMDG 600
QY 601 RNNVENVFINKPQSGTYTIEVOAIVNPVGPQXFSLSAIVN 639
DB 601 RNNVENVFINKPQSGTYTIEVOAIVNPVGPQXFSLSAIVN 639

RESULT 8
AA17088
ID AA17088 standard; protein; 640 AA.
XX
XX AA17088;
AC
AC 20-MAR-2003 (revised)
DT 21-JUN-1999 (first entry)
XX
XX An alkaline protease sequence from Bacillus species.
DE
DE Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW washing composition; oxidizing agent.
XX
XX Bacillus sp.
OS
OS Key Location/Qualifiers
FH Misc-difference 1..640
FT /note="all residues indicated as Xaa are arbitrary amino
FT acids"
XX
XX MO9918218-A1.
XX
XX 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX (KAOS ) KAO CORP.
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
PI Shikata S, Nomura M;
DR WPI; 1999-287736/27.
XX N-PSDB; AAX37278.
XX
XX Alkali protease from Bacillus used in washing powders.
XX
XX Claim 3; Page 50-53; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease of the
CC invention. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 640 AA;

Query Match      98.3%; Score 3007; DB 2; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.5e-235;
```

```

XX PF 09-JUN-1998; 98WC-US012005.
XX PR 12-JUN-1997; 97US-00873479.
XX RA (NOVO ) NOVO NORDDISK BIOTEC INC.
XX PI Sloma A, Christlanson L;
XX DR WPI; 1999-080908/07.
XX DR N-PSDB; AAV82382.
XX PT Novel protease from Bacillus subtilis LC20 - useful in laundry and
XX PS dishwashing detergents and for leather processing.
XX PS Claim 7; Page 53-54; 77pp; English.
XX
XX This is the amino acid sequence of a novel protease of Bacillus sp. JPI707
XX (NCIB 12513), as deduced from the nucleotide sequence of an isolated gene
XX (see AAV82382). The entire protein, including the signal peptide and
XX propeptide region, has 77% identity to alkaline protease Y (see AAV89548)
XX from Bacillus. The invention provides vectors, recombinant host cells and
XX methods for the recombinant production of the protease. The protease is
XX used in laundry and dishwashing detergents, for institutional and
XX industrial cleaning, and for leather processing, as well as for
XX delecterizing and enhancing the degree of hydrolysis of protein
XX hydrolysates, for flavour development through hydrolysis of proteins,
XX degradation of undesired peptides and in enzymatic synthesis of peptides
XX It has enhanced stability towards oxidation under alkaline conditions,
XX e.g. towards bleaching agents of the peroxy type. The invention also
XX provides mutant cells in which the protease activity is diminished. Such
XX cells can be used for the production of heterologous recombinant proteins.
XX
XX Sequence 641 AA.
XX
XX Query Match 89.0%; Score 2723; DB 2; Length 641;
XX Best Local Similarity 82.9%; Pred. No. 3e-212;
XX Matches 532; Conservative 40; Mismatches 66; Indels 4; Gaps 2;
XX
Oy 1 MRK--KKYELSTSAALSTVALKNPSAGAKRFDLDFGCIQTDDXKFSKQXGTGA 57
Db 1 MRKKGSRVFLSTLVAALLSSVALSSPBTIGANNFELDFGFIETLLEKATQKGTGK 60
Oy 58 AALVSESENVKTLKGLKKLSTVPANNKLHIXQFNGPILTEETKOKLXTGAKLIDY.PDY 117
Db 61 ASFLVSESENVKLPKSLQKGLGVPPADNKLIVQFPGPILTEETQQLQLETKAKLIDY.PDY 120
Oy 118 AYIVYEGDVKSXXXIXEHVESVEEYLPYXIIDPOLFTGKASLKVKXALDTQKXKEVQ 177
Db 121 AYIVERDGVKAVTMAIAHLSVEEYLPYKIDQLPFRGASSELVEYVALDKKQRSKEVR 180
Oy 178 LRGLGKIAQXXKSNDVXYITTAPEYKVMNDVAKGVKADVAAQSSYGLYGQCIYAAVADTG 237
Db 181 LRGLGKIAQYATATNDLVLTTPPEYKVEVINDVAKGVKADVAAQNNFGYGGQCIYAAVADTG 240
Oy 238 LDRGRDSSMHEAFEGKTLAYALGRTNNADVDTNGHGHVAGSLGKNTKGMAPANL 297
Db 241 LDRGRDSSMHEAFEGKTLAYALGRTNNADVDTNGHGHVAGSVLGN-ATTKGAAPANL 299
Oy 298 VFOSINDSXGGLGGLPSNLTQTLFSQAKSAGARIHTNSGCAVNGAYTTDSRNVDDYVRKN 357
Db 300 VFOSINDSGGGLGGLGGLPANLTQLFSGAYAGARIHTNSGCAVNGAYTTDSRNVDDYVRKN 358
Oy 358 DMTTLTAAAGNEKPNGGTISAPGTAAKALTVGATENLRFSGSYADNINHTVAQSSRGPTK 417
Db 360 DMTTLTAAAGNEKPGSGTISAPGTAAKALTVGATENLRFSGSYADNINHTVAQSSRGPTK 419
Oy 418 DGRKFDVWAPGXYILSARSSLAPDSSFWAHDSRYVAMGTSWATPIVAGVAAQLREHF 477
Db 420 DGRKFDVWAPGXYILSARSSLAPDSSFWAHDSRYVAMGTSWATPIVAGVAAQLREHF 479
Oy 478 VKNRGITTPPSLLKALTAGADKGLGYPNGQCGRVTLDKSLNVAVNESSXLSTPSQK 537

```



```

Db      480 VKNRGVTPEKSLKAALLIAGADVGCGFENGQCKGRVTLDKSLNAPVNETSPLSTSQK 539
Qy      538 AITXFTATRGKPKLSLWMSDABASTYAVTLVNDLDLVTAPEGTYVYGNDPFXEXXXN 597
Db      540 ATYSPTAOGKPKLSLWMSDABGSTTASLTIVNDLDVITAPEGNTKYVGNDFATAYDNK 599
Qy      598 WQGRNNVENVFINKPOSCTTIIVQAINPVFGQKTSIAIVN 639
Db      600 WDGRNNVENVFINKPOSCTTYVEVQAIXNPVSPOTFSIAIVH 641

RESULT 10
ID      AAM89548 standard; protein; 636 AA.
AC      AAM89548;
DT      12-APR-1999 (first entry)
DB      Bacillus sp. alkaline protease Y.
DS      Alkaline protease Y; detergent; surfactant; leather processing;
KM      debittering; flavour.
OS      Bacillus sp.
PN      WO9656927-A2.
PD      17-DEC-1998.
PR      09-JUN-1998; 98WO-US012005.
PX      12-JUN-1997; 97US-00873479.
PA      (NOVO ) NOVO NORDISK BIOTECH INC.
PI      Sloma A, Christianson L;
PP      WPI; 1999-080908/07.
PT      Novel protease from Bacillus subtilis LC20 - useful in laundry and
PS      dishwashing detergents and for leather processing.
        Claim 3; Page 55-56; 77pp; English.
XX      This is the amino acid sequence of a Bacillus sp. alkaline protease X
XX      that is said to have good alkali and surfactant resistance and improved
XX      detergentcy. It shows 77% identity to a newly isolated protease (see
XX      AM89547) of Bacillus sp. JFI170 (NCIB 12513). The invention provides
XX      vectors, recombinant host cells and methods for the recombinant
XX      production of such proteases. The protease are used in laundry and
XX      dishwashing detergents, for institutional and industrial cleaning, and
XX      for leather processing, as well as for debittering and enhancing the
XX      degree of hydrolysis of protein hydrolysates, for flavour development
XX      through hydrolysis of proteins. Degradation of undesired peptides and in
XX      enzymatic synthesis of peptides. They have enhanced stability towards
XX      oxidation under alkaline conditions, e.g. towards bleaching agents of the
XX      peroxy type. The invention also provides mutant cells in which the
XX      protease activity is diminished. Such cells can be used for the
XX      production of heterologous recombinant proteins
XX      Sequence 636 AA;
SQ

Query Match          78.7%; Score 2408; DB 2; Length 636;
Best Local Similarity 73.2%; Pred. No.1,2e-186;
Matches 467; Conservative 67; Mismatches 98; Indels 6; Gaps 5;
Qy      3 KKKVFLSVLSAAAIISTVALKNPSAGKARXPDLDFKGIGTTDDXXGFKOQTGAARLV 62
Db      4 KKRVLVSVAASAIIASVWSPTSGLA--DFQNVNGVK-SLENASLVKPISSGEASFLV 60
Qy      63 ESENVKLAKGLKKGLDETVPANNKHINQN-GPILEETKOKILETKGAILDIYPYATIV 121

```

Db	61	DTENINIPKGIQKLEAVQKDNELVIVQPTPGPISEEEKKGLSLGLSVILDDVYPDAFIY	120		
Qy	122	ETEGDVXSKXXXIHEHVESVEPYLPXXIIDPOLFTGASLTVAXALDTQXNKEVOLRG	181		
Db	121	QVSG-ATKNISTLHSEVENQPLPYKIDPELLTGASQLVQAVIINTGHEKNMKPTGL	179		
Qy	182	EXIAQXXXNDVXYITTAKEPYVMNDVARGIVKADVAQSSYGLYGGQIVAVADTGLDTG	241		
Db	180	DEIVQYAAANNDDVLYISPKREYELMDVANGIVKADVAQNNYGLYGGQIVAVADTGLDTG	239		
Qy	242	RNDSSMHEAFRKKITALLVALGRTNNANDTNNGHTHVAGSVLGNCKTKNKGAPQANLVFQS	301		
Db	240	RNDSSMHEAFRKKITALLVALGRTNNASDPNGGHTHVAGSVLGN-ALNKGMAPOANLVFQS	298		
Qy	302	INDSXGIGLGLPSNQTLPQXKSSAGARIHTNSWGAIVNGAVATTTSSRNVDVYRKDDMTI	361		
Db	299	INDSSGGLGLPLSNLNTLPQAMNAGARLHTNSWGAIVNGAVATTAUSRQVDEYVRNDDTV	358		
Qy	362	LPAAGNEXPNGGTISAPGTAKNAITVGATENTLRPSFGSYADNINHVAQPSRSRGPTQDGI	421		
Db	359	LPAAGNEXPNGGTISAPGTAKNAITVGATENTLRPSFGSLADPNHIAQPSRSRGATDGI	418		
Qy	422	KPDVVAPEGTXIISANSSLAPODSEFMANHDSKTYAVGCTSMATPIYAGNVAQLREHFVKR	481		
Db	419	KPDVVAPEGTXIISARSSLAPODSEFMANVYSKYAVMGCTSMATPIYAGNVAQLREHFVKR	478		
Qy	482	GITPPESLIKKALIGAADXGLGYGPNQMGWRVTLKDSLNAVYVNESKSLSTSOKATYX	541		
Db	479	GITPPESLIKKALIGATDVGIGYSGDQMGWRVTLKDSLNAVYVNEATALLATGQATYS	538		
Qy	542	FTATAGPELKI SLVMSDAPASTTASVTLVNDLIVITAPNGTXKYVGNDFXKEXXNMGR	601		
Db	539	FOAQGRDLKISLVMTDAPGSTTASVTLVNDLIVITAPNGKYVGNDFPSAIVDNNMGR	598		
Qy	602	NNVENVFNNPQSGTYTIEVQANVPVGPQFSLAIYN	639		
Db	599	NNVENVFNNPQSGTYTIEVQANVPVGPQFSLAIYH	636		
RESULT 11					
AAMS0080					
ID	AAMS0080 standard; protein; 434 AA.				
XX	AC	AAMS0080;			
XX	DT	12-AUG-2002 (first entry)			
XX	DE	Bacillus sp KSM-KP43 alkaline protease protein fragment.			
XX	KW	Alkaline protease; detergent; laundry; bleaching; dishwasher.			
XX	OS	Bacillus sp.			
XX	FN	EP1209233-A2.			
XX	PD	29-MAY-2002.			
XX	PE	22-NOV-2001; 2001EP-00127851.			
XX	PR	22-NOV-2000; 2000JP-00355166.			
XX	PA	12-APR-2001; 2001JP-00114048.			
XX	PA	(KAOS) KAO CORP.			
PI	Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N,				
PI	Okuda M, Saeki K;				
XX	WPI; 2002-437518/47.				
XX	New modified alkaline proteases useful in detergent compositions.				
XX	Claim 1, Page 10-11; 25pp; English.				

CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency *
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease KP43 from
 CC *Bacillus* sp strain KSM-KP43 which is used to create the modified protease
 CC represented in AAM50090

XX
 CC
 SQ Sequence 434 AA;

Query Match 70.4%; Score 2155; DB 5; Length 434;
 Best Local Similarity 96.3%; Pred. No. 2.5e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 206 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITALYALGRTN 265
 DB 1 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITALYALGRTN 60
 QY 266 NNDTNGHGTHTVAGSVLGNKXTNKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFSGQAYS 325
 DB 61 NNDTNGHGTHTVAGSVLGNKXTNKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFSGQAYS 120
 QY 326 AGARHTNSWGAUVNGAYTTDSRNVDVYRKNDMTLLFAAGNEXPGGTISAPGTAKNAI 385
 DB 121 AGARHTNSWGAUVNGAYTTDSRNVDVYRKNDMTLLFAAGNEXPGGTISAPGTAKNAI 180
 QY 386 TVGATENTLRPSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSF 445
 DB 181 TVGATENTLRPSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSF 240
 QY 446 WANHDSKYAVMGSTMAPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAGADGIGY 505
 DB 241 WANHDSKYAVMGSTMAPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAGADGIGY 300
 QY 506 PNGNGMGKRVTLDKSLNVAAYVNESSSLSTSOAKATYXFTATAGKPLKISLWSDAPASTTA 565
 DB 301 PNGNGMGKRVTLDKSLNVAAYVNESSSLSTSOAKATYXFTATAGKPLKISLWSDAPASTTA 360
 QY 566 SYTLVNDLDELVTAPNGTXYVGNDFPKXPKXKXMDGRNNVENVFINKPQSGTYTIEVOAYN 625
 DB 361 SYTLVNDLDELVTAPNGTXYVGNDFPKXPKXKXMDGRNNVENVFINKPQSGTYTIEVOAYN 420
 QY 626 VPVGPQXPSLAIYN 639
 DB 421 VPVGPQXPSLAIYN 434

RESULT 12
 AAM50081

ID AAM50081 standard; proteain; 434 AA.

XX AAM50081;
 XX
 AC
 XX
 DT 12-AUG-2002 (first entry)
 XX
 DB *Bacillus* sp KSM-KP9860 alkaline protease proteain fragment.
 XX
 KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS *Bacillus* sp.
 XX
 PN EP1209233-A2.
 XX
 PD 29-MAY-2002.
 XX
 PF 22-NOV-2001; 2001BP-00127851.
 XX
 PR 22-NOV-2000; 2000JP-00355166.
 XX
 PR 12-APR-2001; 2001JP-00114048.
 XX
 PA (KAOS) KAO CORP.
 XX
 PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

PI Okuda M, Saeki K;
 XX
 DR WPI; 2002-437518/47.
 XX
 PT New modified alkaline proteases useful in detergent compositions.
 XX
 PS Claim 5; Page 12-13; 25pp; English.

CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency *
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease KP9860 from
 CC *Bacillus* sp strain KSM-KP9860 described in the method of the invention
 CC
 XX
 SQ Sequence 434 AA;

Query Match 70.4%; Score 2155; DB 5; Length 434;
 Best Local Similarity 96.3%; Pred. No. 2.5e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 206 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITALYALGRTN 265
 DB 1 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITALYALGRTN 60
 QY 266 NNDTNGHGTHTVAGSVLGNKXTNKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFSGQAYS 325
 DB 61 NNDTNGHGTHTVAGSVLGNKXTNKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFSGQAYS 120
 QY 326 AGARHTNSWGAUVNGAYTTDSRNVDVYRKNDMTLLFAAGNEXPGGTISAPGTAKNAI 385
 DB 121 AGARHTNSWGAUVNGAYTTDSRNVDVYRKNDMTLLFAAGNEXPGGTISAPGTAKNAI 180
 QY 386 TVGATENTLRPSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSF 445
 DB 181 TVGATENTLRPSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSF 240
 QY 446 WANHDSKYAVMGSTMAPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAGADGIGY 505
 DB 241 WANHDSKYAVMGSTMAPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAGADGIGY 300
 QY 506 PNGNGMGKRVTLDKSLNVAAYVNESSSLSTSOAKATYXFTATAGKPLKISLWSDAPASTTA 565
 DB 301 PNGNGMGKRVTLDKSLNVAAYVNESSSLSTSOAKATYXFTATAGKPLKISLWSDAPASTTA 360
 QY 566 SYTLVNDLDELVTAPNGTXYVGNDFPKXPKXKXMDGRNNVENVFINKPQSGTYTIEVOAYN 625
 DB 361 SYTLVNDLDELVTAPNGTXYVGNDFPKXPKXKXMDGRNNVENVFINKPQSGTYTIEVOAYN 420
 QY 626 VPVGPQXPSLAIYN 639
 DB 421 VPVGPQXPSLAIYN 434

RESULT 13

ID ADY33778 standard; proteain; 434 AA.

XX ADY33778;
 XX
 AC
 XX
 DT 05-MAY-2005 (first entry)
 XX
 DB *Bacillus* species alkaline protease.
 XX
 KW mutagenesis; alkaline protease; detergent; laundry; cosmetic; food;
 XX
 OS *Bacillus* sp.
 XX
 PN EP1347044-A2.
 XX
 PD 24-SEP-2003.
 XX
 PI

```

PF 21-MAR-2003; 2003EP-0006472.
XX
XX 22-MAR-2002; 2002JP-00081428.
PR 06-JUN-2002; 2002JP-00165987.
PR 18-OCT-2002; 2002JP-00304230.
PR 18-OCT-2002; 2002JP-00304231.
XX
XX (KAOS ) KAO CORP.
XX
XX Okuda M, Sato H, Saito K, Sumitomo N, Izawa Y, Saeki K;
PI Kobayashi T, Nomura M;
DR WPI; 2003-846540/79.
XX N-PSDB; ADY33779.
XX
XX New alkaline protease having specific amino acid residue at a specific
PT position of its amino acid sequence, useful for producing detergent
PT compositions, laundry detergent, fiber modifiers, leather-treating agents
PT or pipe cleaners.
XX
XX Claim 1; SEQ ID NO 1; 31pp; English.
XX
XX The invention relates to an alkaline protease having a fully defined
CC sequence of 434 amino acids (I) given in the specification, or an amino
CC acid sequence at least 80% homology with (I), where an amino acid residue
CC at position 65, 101, 163, 170, 171, 273, 320, 359 or 387 of (I) is
CC selected from 16 amino acid residues. The amino acid residues at the
CC corresponding positions are selected from: position 65: proline; position
CC 101: asparagine; position 163: histidine, aspartic acid, phenylalanine,
CC lysine, asparagine, serine, isoleucine, leucine, glutamine, threonine or
CC valine; position 170: valine or leucine; position 171: alanine, glutamic
CC acid, glycine or threonine; position 273: isoleucine, glycine or
CC threonine; position 320: phenylalanine, valine, threonine, leucine,
CC isoleucine or glycine; position 359: serine, leucine, valine, isoleucine
CC or glutamic acid; and position 387: alanine, lysine, glutamine, glutamic
CC acid, arginine or histidine. The alkaline protease is useful for the
CC production of a detergent composition, such as laundry detergent, fiber
CC modifiers, leather-treating agents, cosmetic compositions, bath
CC additives, food modifiers and pharmaceutical compositions. The alkaline
CC protease may also be used as bleaching detergent, hard surface cleansing
CC detergent, pipe cleaner, artificial tooth cleaner, and as a sterilizing
CC cleanser for medical tools. The new alkaline protease has a more potent
CC proteolytic capacity, exhibiting excellent detergency for the removal of
CC a complex stain, and has high secretion capacity. This sequence
CC corresponds to the Bacillus sp. alkaline protease.
XX
XX Sequence 434 AA;
SQ
Query Match 70.4%; Score 2155; DB 7; Length 434;
Best Local Similarity 96.3%; Pred. No. 2.5e-166;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 206 NDVARGIVKADVAAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALYALGRTN 265
DB 1 NDVARGIVKADVAAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALYALGRTN 60
QY 266 NANTNGGTHVAGSVLGNKGNTKGMAPQANLVPQSIINDSXGGLGSLNLTOTFESQAXS 325
DB 61 NANTNGGTHVAGSVLGNKGNTKGMAPQANLVPQSIINDSXGGLGSLNLTOTFESQAXS 120
QY 336 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILFPAENGEXNGGTSIAPGTAKNAI 385
DB 121 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILFPAENGEXNGGTSIAPGTAKNAI 180
QY 386 TVGATENTRPFPGSIVADNINHTVAQSSRGPTDGRIRKPDVWAPGTXYIIISARSLAPDSSF 445
DB 181 TVGATENTRPFPGSIVADNINHTVAQSSRGPTDGRIRKPDVWAPGTXYIIISARSLAPDSSF 240
QY 446 WANHDSKTAVMGGTSMATPIVAGNVAOLREHPVKRKGITTPKPSLLKALLINGADIXGIGY 505
DB 241 WANHDSKTAVMGGTSMATPIVAGNVAOLREHPVKRKGITTPKPSLLKALLINGADIXGIGY 300
QY 506 PNGNGMGRVTLTDLKSLNVAAYNBSXLSISOKATYXFTATAGKPLKISLVMSDAPASTTA 565

```

```

DB 301 PNGNGMGRVTLTDLKSLNVAAYNBSXLSISOKATYXFTATAGKPLKISLVMSDAPASTTA 360
QY 566 SVTLVNDLVLITVTPNGTXYVGNDEKXXPKXKNMDGRNNVENVFITNXPGSGTYTIEVQAYN 625
DB 361 SVTLVNDLVLITVTPNGTXYVGNDEFTSPYNDNMGRNNVENVFITNAPSGTYTIEVQAYN 420
QY 626 VPPGPRQKSLATVY 639
DB 421 VPPGPRQKSLATVY 434
RESULT 14
ADZ51758
ID ADZ51758 standard; protein; 434 AA.
XX
XX ADZ51758;
AC
XX 16-JUN-2005 (first entry)
DT
XX
DE Mutant Bacillus sp. alkali protease, SEQ ID No:2.
XX
XX alkali protease; enzyme; surfactant; mutagenesis; muteln.
XX
XX Bacillus sp.
OS Synthetic.
XX
XX JP2003125783-A.
XX
XX 07-MAY-2003.
PD
XX
XX 26-OCT-2001; 2001JP-00329472.
PF
XX
XX 26-OCT-2001; 2001JP-00329472.
PR
XX 26-OCT-2001; 2001JP-00329472.
PA (KAOS ) KAO CORP.
XX
XX WPI; 2003-855669/80.
DR
XX
XX New alkali protease useful for preparing detergents comprises
PT substitution mutations at 251 or 256 position of protease KP43 derived
PT from Bacillus species KSM-KP43.
XX
XX Claim 3; SEQ ID NO 2; 16pp; Japanese.
XX
XX The invention relates to a mutant alkali protease having an amino acid
CC deletion at position(s) 251 or 256 in a fully defined sequence given as
CC SEQ ID No:1 in the specification, or the following amino acid residue
CC Asp, Gly, Ala, Asp, Thr, Ile, Val, Leu or Glu at position 251, or the
CC amino acid residue Lys, Ser Phe, Arg, Tyr, Met, Thr, Ile, Val, Leu, or
CC Glu at position 256. Also described are: (i) an alkali protease having
CC the amino acid sequence of SEQ ID No:1, or having 60% or more homology to
CC SEQ ID No:1, or a deletion of amino acid corresponding to position 251 or
CC 256, (ii) a gene which encodes an alkali protease, (iii) a recombinant
CC vector containing the gene, and (iv) a transformed organism containing
CC the recombinant vector. The alkali protease sequence having 60% or more
CC homology to SEQ ID No:1 is chosen from a fully defined sequence selected
CC from SEQ ID Nos 2-7 as given in the specification. The transformed
CC organism is a microorganism. The mutant alkali protease is useful for
CC preparing detergents. The mutant alkali protease exhibits high resistance
CC against oxidizing agent. The mutant alkali protease has high specific
CC activity. This sequence represents a mutant Bacillus sp. alkali protease.
XX
XX Sequence 434 AA;
SQ
Query Match 70.4%; Score 2155; DB 7; Length 434;
Best Local Similarity 96.3%; Pred. No. 2.5e-166;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 206 NDVARGIVKADVAAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALYALGRTN 265
DB 1 NDVARGIVKADVAAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALYALGRTN 60

```

QY 266 NANDTNGHGTTHVAGSVTLGNGXTNKGMAPQANLVFQSIIMDSXGGLGSLPSNLQTLTLPFOAYS 325
DB 61 NANDTNGHGTTHVAGSVTLGNGXTNKGMAPQANLVFQSIIMDSXGGLGSLPSNLQTLTLPFOAYS 120
QY 326 AGARLHTNSMGAAVNGAYTTTDSRNVDYVRKNDMTLLFAAGNEXPNCGTISAPGTAKNAI 385
DB 121 AGARLHTNSMGAAVNGAYTTTDSRNVDYVRKNDMTLLFAAGNEXPNCGTISAPGTAKNAI 180
QY 386 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPGTYILSARSSLAPDSSF 445
DB 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPGTYILSARSSLAPDSSF 240
QY 446 WANHSKAYMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAADVGLGY 505
DB 241 WANHSKAYMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAADVGLGY 300
QY 506 PNGNGMGKRVTLDKSLNVAAYVNESSXLSTQKATYXFTATAGKPLKISLWSDAPASTTA 565
DB 301 PNGNGMGKRVTLDKSLNVAAYVNESSXLSTQKATYXFTATAGKPLKISLWSDAPASTTA 360
QY 566 SVTLVNDLDELVTAPNGTYVGNDFPKXPKXKXNDGRNNVENVFINKPQSGTYTIEVOAYN 625
DB 361 SVTLVNDLDELVTAPNGTYVGNDFPKXPKXKXNDGRNNVENVFINKPQSGTYTIEVOAYN 420
QY 626 VPVGPOXPSLAIYN 639
DB 421 VPVGPOXPSLAIYN 434

RESULT 15
AD251757
ID AD251757 standard; protein; 434 AA.
XX
AC AD251757;
XX
DT 16-JUN-2005 (first entry)
XX
DB Wild-type Bacillus sp. alkali protease.
XX
KW alkali protease; enzyme; surfactant; mutagenesis.
XX
OS Bacillus sp.
XX
PN JP2003125783-A.
XX
PD 07-MAY-2003.
XX
PF 26-OCT-2001; 2P01JP-00329472.
XX
PR 26-OCT-2001; 2001JP-00329472.
XX
PA (KAOs) KAO CORP.
XX
DR WPI; 2003-856569/80.
XX
PT New alkali protease useful for preparing detergents comprises
PT substitution mutations at 251 or 256 position of protease KP43 derived
PT from Bacillus species KSM-KP43.
XX
PS Claim 1; SEQ ID NO 1; 16pp; Japanese.
XX
CC The invention relates to a mutant alkali protease having an amino acid
CC deletion at position(s) 251 or 256 in a fully defined sequence given as
CC SEQ ID No.1 in the specification, or the following amino acid residue
CC Asp, Gly, Ala, Asp, Thr, Ile, Val, Leu or Glu at position 251, or the
CC amino acid residue Lys, Ser, Phe, Arg, Tyr, Met, Thr, Ile, Val, Leu, or
CC Glu at position 256. Also described are: (i) an alkali protease having
CC the amino acid sequence of SEQ ID No.1, or having 60% or more homology to
CC SEQ ID No.1, or a deletion of amino acid corresponding to position 251 or
CC 256, (ii) a gene which encodes an alkali protease, (iii) a recombinant
CC vector containing the gene, and (iv) a transformed organism containing
CC the recombinant vector. The alkali protease sequence having 60% or more
CC homology to SEQ ID No.1 is chosen from a fully defined sequence selected

CC from SEQ ID Nos 2-7 as given in the specification. The transformed
CC organism is a microorganism. The mutant alkali protease is useful for
CC preparing detergents. The mutant alkali protease exhibits high resistance
CC against oxidizing agent. The mutant alkali protease has high specific
CC activity. This sequence represents wild-type Bacillus sp. alkali
CC protease.
XX
SQ Sequence 434 AA;
Query Match 70.4%; Score 2155; DB 7; Length 434;
Best Local Similarity 96.3%; Pred. No. 2.5e-166;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 206 NDVARGIVADVAQSSYGLYGQGIYVAVDTGLDTRNDSSMHEAFRGKITLALYALGRYN 265
DB 1 NDVARGIVADVAQSSYGLYGQGIYVAVDTGLDTRNDSSMHEAFRGKITLALYALGRYN 60
QY 266 NANDTNGHGTTHVAGSVTLGNGXTNKGMAPQANLVFQSIIMDSXGGLGSLPSNLQTLTLPFOAYS 325
DB 61 NANDTNGHGTTHVAGSVTLGNGXTNKGMAPQANLVFQSIIMDSXGGLGSLPSNLQTLTLPFOAYS 120
QY 326 AGARLHTNSMGAAVNGAYTTTDSRNVDYVRKNDMTLLFAAGNEXPNCGTISAPGTAKNAI 385
DB 121 AGARLHTNSMGAAVNGAYTTTDSRNVDYVRKNDMTLLFAAGNEXPNCGTISAPGTAKNAI 180
QY 386 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPGTYILSARSSLAPDSSF 445
DB 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPGTYILSARSSLAPDSSF 240
QY 446 WANHSKAYMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAADVGLGY 505
DB 241 WANHSKAYMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAADVGLGY 300
QY 506 PNGNGMGKRVTLDKSLNVAAYVNESSXLSTQKATYXFTATAGKPLKISLWSDAPASTTA 565
DB 301 PNGNGMGKRVTLDKSLNVAAYVNESSXLSTQKATYXFTATAGKPLKISLWSDAPASTTA 360
QY 566 SVTLVNDLDELVTAPNGTYVGNDFPKXPKXKXNDGRNNVENVFINKPQSGTYTIEVOAYN 625
DB 361 SVTLVNDLDELVTAPNGTYVGNDFPKXPKXKXNDGRNNVENVFINKPQSGTYTIEVOAYN 420
QY 626 VPVGPOXPSLAIYN 639
DB 421 VPVGPOXPSLAIYN 434

Search completed: April 7, 2006, 17:37:30
Job time : 140.392 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 17:37:51 ; Search time 27.4785 Seconds

(without alignment)

Title: US-10-784-870-1

Perfect score: 3059

Sequence: 1 MRKKVPLSLVSAALISLV.....EVOGAVNPVGPQPSLAIYN 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518.5	16.9	1743	2	T18279
2	477.5	15.6	1905	2	T18267
3	350.5	11.5	444	2	B83891
4	319.5	10.4	442	2	A69587
5	297.5	9.7	1398	2	T28159
6	286	9.3	806	2	A41341
7	279.5	9.1	419	1	S25835
8	276.5	9.0	799	1	G83753
9	274.5	9.0	420	1	S23407
10	273.5	8.9	580	2	S11890
11	268.5	8.8	715	2	JC4908
12	262	8.6	894	2	P68730
13	260.5	8.5	1345	2	T28090
14	257.5	8.4	378	2	A33973
15	255.5	8.4	381	2	JH0778
16	255	8.3	513	1	A35742
17	253.5	8.3	381	1	SUBSI
18	253	8.3	382	1	SUBSN
19	251.5	8.2	381	1	SUBSS
20	250.5	8.2	380	2	A49778
21	250.5	8.2	602	2	JC4576
22	250	8.2	382	2	T39780
23	249.5	8.2	379	1	SUBSC
24	249.5	8.2	757	2	C84120
25	248	8.1	627	2	D75393
26	247.5	8.1	381	2	JQ1487
27	246.5	8.1	402	1	JU0332
28	242	7.9	1331	2	A72647
29	239.5	7.8	401	2	I39974

30	239.5	7.8	534	1	JS0173	alkaline proteinas
31	235	7.7	519	2	S71451	haloalysin R4 (EC 3
32	232	7.6	488	2	A11930	proteinas (import
33	231.5	7.6	1167	1	A35066	stretococcal Csa
34	230	7.5	384	2	JC4802	alkaline proteinas
35	223.5	7.3	409	1	S32905	serine proteinas
36	223	7.3	321	1	S27501	alkaline proteinas
37	220.5	7.2	535	2	B82358	alkaline serine pr
38	219.5	7.2	613	2	S75976	hypothetical prote
39	218.5	7.1	272	2	A23624	subtilisin (EC 3.4
40	218.5	7.1	275	2	JC1085	serine proteinas
41	218.5	7.1	533	1	JU0146	subtilisin (EC 3.4
42	218	7.1	361	2	G83756	subtilisin-type al
43	216.5	7.1	361	2	A48373	high-alkaline seri
44	216.5	7.1	525	2	G84406	haloalysin (impor
45	216.5	7.1	645	1	SUBSMP	serine proteinas

ALIGNMENTS

RESULT 1

multidrug resistance transport protein - alime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18279
R:Shaulsky, G.; Loomis, W.P.
A:Submitted to the EMBL Data Library, June 1996
A:Reference number: Z18855
A:Accession: T18279
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1743 <SHA>
A:Cross-references: UNIPROT:Q23868; UNIPARC:UPI000013687D; EMBL:U60086; NID:G1399914; PI
A:Gene: tagc

Query Match 16.9%; Score 518.5; DB 2; Length 1743;
Best Local Similarity 24.2%; Pred. No. 2.8e-28;
Matches 194; Conservative 114; Mismatches 213; Indels 281; Gaps 33;

QY	67	HIXQFNGPIERTKXKL-----EXTGAKILDYIPDVAIVY---EYEDVNSX	130
DB	115	YIVQFKRINDTRQLKPLIGTDIVDEQYSHIVYIPDSEFLVMTQSGVLLSS	174
QY	131	XXXIHEVESVEPY-----LPYXIDPOLFTKGSXLVK--AXALDT--KQXN	173
DB	175	KEMVWIGEFERSNKHLMVNEKSTGLPVYI---LSDSTNSLQKMENTLNSILKSYN	230
QY	174	KEVQLRGI-----EKIAQXXXNDVXYITAKPEYK	203
DB	231	SKVKLTLLNQKKLKSIVCNDESPSPSCSLINSEKLVOIWISQESBSN---FIRSEKQ	287
QY	204	VNADVARGIV-----KADVAQSSYGLYGGGQVVAVADPTLDGR--NDS----	245
DB	288	TANRLSPFVPEGTQDTLVNNDKRDIP---LRGQQLISTADTGLDSCHPFSDSKYP	342
QY	246	---SMBAPFGKITALVALGRNNANDTNGGTHVAGSVLG-----NGXNKGAPDA	295
DB	343	PLNSVNLNHR-KVTVYITTSDDSKYDGHGTHICGSAAGTPBSSVNISPSGLADH	401
QY	236	NLVFQSIQSGGLGL--PSNLQTLFQOAXSAGARIHTNSGWA-----AVNGAYTTDSRN	349
DB	402	KIAF---FDLASGSSSLPSPDLKQLYQPLVDAGARVHCDGSGVSGVSSTAS	458
QY	350	VDDYARKN-DNTIIFLAENKXPNGSTIS--ABGTAKNAITYGATNMLR-----PSGGS	399
DB	459	IDDFLFTPHDFIILAAQN---NEGYLSLTQSTAKNVITVGANQOTIHENYLTDDPNYIN	515
QY	400	YADNI-----	404
DB	516	YQSSVDINQELICPDSRYCMTTACCLBSNATTLGLASCCPTLLRKSVIDAANTQPLLY	575

QY 405 --NHVAQSSRGPTKGRIRKPDVMAFGTIIISARSSLA-----PDSSFWAHDSKRYAM 456
DB 576 NENNICSSRSSKGTTHDGRKPKALVAPGEYITARSANGATTDOCGGSI--PTNMLA-I 633
QY 457 GGTSMATPIVAGNVAQLREH-----FYKNRGITPKPSLLKRAALJAGA----- 498
DB 634 SGTSMATSPAAATAATILRLQYLDVGYPTGSIYESKNLQPTGSLLRALMINNAQLNGTFQ 693
QY 499 --ADGGLGYPNGN-----QGKGRVTLDKSLNVAYNSS----- 530
DB 694 LITSSSIYPSNQVFNPAQSLVQGWGAIHRSMNMLHVNNNNNSNNNKTSDGITKFDGI 753
QY 531 -----XLISTOKATYXFT-----ATAGKPLK---ISLWMSDAPAST 564
DB 754 GGLDLRLKPNQWKESLSTGQNTSYCTFYKPSSSSSNSGNNIPRVVATLWTDPSPYAG 813
QY 565 ASVTLVNDLVI-----TAPNGTYVGNDFXXPKXXWDGRNNEVNFIXP 612
DB 814 AKFNLVNNLDMITMYRDNSTIFYSNCGSSFLG-----LAPQDTLNVGIVHNP 867
QY 613 QSGTITIEVQAVNPVPGQXFS 634
DB 868 EEMTYRFVAVGATNVPWGQNF 889

RESULT 2

T18267
multidrug resistance protein - slime mold (Dictyostelium discoideum)
C/Species: Dictyostelium discoideum
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T18267
R.Shaulsky, G.J. Kuapa, A.J. Loomis, W.F.
submitted to the EMBL Data Library, January 1995
A/Description: An MDR transporter/serine protease gene is required for prestalk speciali
A/Reference number: Z18850
A/Accession: T18267
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1905 <SHA>
A/Cross-references: UNIPROT:P54683; UNIPARC:UPI000013687C; EMBL:U20432; NID:g664839; PID
C/Genetics:
A/Gene: tagB

Query Match 15.6%; Score 477.5; DB 2; Length 1905;
Best Local Similarity 24.5%; Pred. No. 2.5e-25;
Matches 193; Conservative 98; Mismatches 237; Indels 259; Gaps 31;
QY 87 HIXQFNGPILEBTKQXLE---XTGAKILDYIPDYAVIYVY-----EGDYXS 129
DB 172 YIVQPKDINDETRQLEKFLGTDTITLKQPFKSHIVHYIPIHDSFLVMTKESVLS 231
QY 130 XXXXIHEVESVPEY-----LPXYI---DPQLFTGASKYL-----V 162
DB 232 SKEMSWIGEHFSPKNIKILNYHEKSIQYPVYIILSGTNSLIQWMENTLNSITSYNSKV 291
QY 163 KAXALDTHKO-----XNKEVQLRGIEIXIAQXXX--NDVXYITAKEKYKM 205
DB 292 KTLTLNQKLSIYCNDESPSSSSCSLIGSKIIYKMISESSBNYIERSEKLOTA 351
QY 206 NDVARGIV-----KADVAGSYGLYGQGIIVADVADTGLDTGR---NDS----- 245
DB 352 NRLSPTVIFGTRKDLVNNDRIDIP-----LRGKQGLISLADTGLDGSCFPSSDKYPIPF 406
QY 246 -SMHEAFKGTITLALARTNNANDTNGHGVAGSVLGNKXTN-----KMAPQANL 297
DB 407 NQVVENHKKVYV---YITYHNDIEDYVNGHGVHVCSSAAGTPEDSSMAISFSGLATDAKI 463
QY 298 VFGSMDXGIGLPSNLQTLFSGAKSAGARIHTNSGA-----AVNAGYTTDSNVDY 353
DB 464 AFYD-LSSGSSBPPEBDSQYKPLVDAGARVHDSGWSVLSQGYGYGSDADAGIDAF 522
QY 354 VRK-NDMTILPAAGNEXPNGTISAPGTAKNAITVGATENLRPFGS-----YADNI- 404

DB 523 LYHEPESITLAAAGNN-ELFASLAAQTAKAATVGAEGTAAHVNVSDALEYPSDAN 581
QY 405 -----NHVAQF 410
DB 582 FQRPCLFDKRYCNNTAKCCSEVENVKGLQLCCPASIKONASDFTTQPPQFENNMSGF 641
QY 411 SSRPTKGRIRKPDVMAFGTIIISARSS-----LAPSSFWAHDSKRYAMGTSMATP 464
DB 642 SSKPTHDGRKPKDIPVAPGEYITARSNGENSTDOCGGSI--DNANLMSISGSMATP 699
QY 465 IVAGNVAQLREH-----YKNRGITPKPSLLKRAALI----- 495
DB 700 LATATITLRQYLDVGYPTGSEVBENKLEPTGSLLRALMINNAQLNGTYFWASSTNP 759
QY 496 AGAADXGLGYPNQGWGRVTLDKSLNVAYNSS-----XLISTOKAT- 539
DB 760 SNAIFEQINGANLIGWALRMN---NWLKYSNPPTPPSRWIGIGLGRKQKATWKKED 816
QY 540 -----YXFT-----ATAGKP-LKISLWMSDAPASTASVTLVNDL- 575
DB 817 SLSSGLKSYCTFYKPSSSSSSGGGGCTPRIVATLWTDPSPYSGAKFNLVNNLDDL 876
QY 576 -----VITAPN---GTYYVGNDFXXPKXXWDGRNNEVNFIXPGGTYTEVQAVNP 627
DB 877 NSDDDSIITIGNSGGSLQAPAKVAQ-----DTLANNVEGIIINPTKAMNYKFTIAGTNV 931
QY 628 VGPQXFS 634
DB 932 IGPQKFS 938

RESULT 3

B83891
intracellular alkaline serine proteinase aprX [imported] - Bacillus halodurans (strain C
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: B83891
R.Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saeaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: B83891
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-444 <STO>
A/Cross-references: UNIPROT:Q9KBU7; UNIPARC:UPI00000C3D43; GB:AP001513; GB:BA000004; NID
A/Experimental source: strain C-125
C/Genetics:
A/Gene: aprX

Query Match 11.5%; Score 350.5; DB 2; Length 444;
Best Local Similarity 30.2%; Pred. No. 3.2e-17;
Matches 114; Conservative 54; Mismatches 127; Indels 83; Gaps 17;
QY 182 EXIAQXXXNDVXYITAKPEYKKNVDVARGIVRA-DVAGSSYGLYGQGIIVADVADTGLDT 240
DB 100 ESLQEMLVCKDIRKIYLNREVNHALDTRVESAQAFVIRNGETTLTGKOVTAVIDGI-- 157
QY 241 GRNDSMHEAFRGKITALY-ALGRTNNANDTNGHGVAGSVLGNKXTN-----KMAPQA 295
DB 158 -----YHEDLEGRIKAFVDFVNGRREPYDNGHGTCHCADDAAAGNASSGQYRGAPAEA 212
QY 296 NLVFGSMDXGIGLPSNLQTLFSGAKSAGARIHTNSGAALVNGAYTTDSRNV----- 350
DB 213 NVIGVKTIANRO-GWGSLESIMQGV-----EWCIQYNEHPPDDPIHISMSL 257
QY 351 -----DDVYV-----KNMTILPAAGNEXPNGTISAPGTAKNAITVGATEN 392
DB 258 GGOALPYENBOEDPMVAVIIEBAMNAGITVCVAAGNSGPDQTIASPVSEKVTIVGALDD 317
QY 393 LRPEFGYADININVAQFSSRGPTKGRIRKPDVMAFGTIIISARSSLAPDSSF-----WA 447

Db 318 -RDITDREDD---VAFSSRGPTIYGKPEKPDILAPGVNIVLSRS---PNSFYDKLQKGS 370
QY 448 NNDSTKAVYGGTSMATPIYAGNVAGQREHFVKRKGITTPPSLKALALIGA---AD----- 500
Db 371 RVGSHYTMVSGTSMATPVCGVVALMLQH---EPNLTPE--VKRLMEGTDRMADRDN 425
QY 501 -XGLGY-----PENGQ 510
Db 426 VYGAGTISABGALPNSBE 443

RESULT 4

intracellular alkaline serine proteinase aprx - Bacillus subtilis
A69587
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: A69587
R/Kunitz, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beretz
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fume, S.; Galizzi, A.; Gallen
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Lardinois,
Koecker, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maseel
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, B.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A/Authors: Scheich, S.; Schreier, R.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Seron
Akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Tepstra, P.; Tognoni, A.; Tosiato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumelein, B.; Yoshikawa, H.; Darchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; M01D:9804033; PMID:9384377
A/Accession: A69587
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-442 <KUN>
A/Cross-references: UNIPROT:O11788; UNIPARC:UPI000006043P; GB:Z99113; GB:AL009126; NID:9
A/Experimental source: strain 168
C/Genetics:
A/Genes: aprX
F:146-398/Domain: subtilisin homology <SBT>

Query Match 10.4%; Score 319.5; DB 2; Length 442;
Best Local Similarity 30.0%; Pred. No. 4.9e-15;
Matches 106; Conservative 46; Mismatches 108; Indels 93; Gaps 15;
QY 201 EYKMNADVARGIVKA-DVAQSSVGLYGQGIIVADVADTGIDGRNDSMHEAFRGKTALY 259
Db 120 EYKALIDRTAESHAKAVNRNGDTLTKGVYAAVVDGI-----YHPDLBGR1---- 168
QY 260 ALGRTNNAN-----DTNGHGHVAGSVLNGXTN-----KGMAPQANLVFQSIIMDSXGL 309
Db 169 -IGPADMVVQKTEPRDDNGHGHGHCADVASSGSGQYRPAPEANLIGVAVLNQGS- 226
QY 310 GGLPSNLQTLFQSAKSAGARIHNSGAAVN-----GAYTDSBRVVD 352
Db 227 GTLADIIEGV-----EKCIOYNEDNDEPIDIMSMISGIGDALVYDHQEDP 272
QY 353 YVRKND-----MTILFAAGNEKPNGGTISAPGTAKNATITVGAETELRSPFSYADNIN 406
Db 273 LVRAVEAMASAGIVCVAGNAGSGPDSQITIASPEVSEKVTITVGLADNNTA-----SSDDPT 328
QY 407 VAFSSRGPTKDGRIKPDVMAFGTILSARSSILAPDSF-----VANHDSKYAVMGCTSM 461
Db 329 VAFSSRGPTVYGKPEKPDILAPGVNIVLSRS---PNSYIDKLQKSRVSGQFTMSGISM 385
QY 462 ATPYVAGNVAGQREHFVKRKGITTPK--PSLKALALIGAADGGLGYPNGNQW 512
Db 386 ATPICAGIALALIQ---QNPDLTPBEVKELK-----NGTIDKX 420

RESULT 5

T28159
pyrolysin (EC 3.4.-.-) - Pyrococcus furiosus
C/Species: Pyrococcus furiosus
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T28159
R/Voohorst, W.G.B.; Eggen, R.I.L.; Geerling, A.C.M.; Plateeuw, C.; Sieren, R.J.; Vos,
J. Biol. Chem. 271, 20426-20431, 1996
A/Title: Isolation and characterization of the hyperthermostable serine protease, pyroly
A/Reference number: Z20481; M01D:96355370; PMID:8702780
A/Accession: T28159
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1398 <VOO>
A/Cross-references: UNIPROT:P72186; UNIPARC:UPI000016FDS9; EMBL:U55835; NID:91556462; P
A/Experimental source: DSM3638
C/Genetics:
A/Genes: p18
C/Keywords: hydrolase; serine proteinase

Query Match 9.7%; Score 297.5; DB 2; Length 1398;
Best Local Similarity 22.9%; Pred. No. 8.5e-13;
Matches 173; Conservative 76; Mismatches 248; Indels 257; Gaps 31;
QY 96 LRETKQXLEKTKAKLID-----YIPYATVYRE-----GDYKXXXXXB 135
Db 77 LBSATLEKLGABILDENRVLNMLVKIKPEKXVLSLSLEKAMLRVGLSPPIVE 136
QY 136 -HVESVEPYLPYXIDPOLFT-----KGASXIVKAXALDTPKQXKREY 176
Db 137 KVKTKREP-----SLBPKVNSTWVYINMLQFIQEGYQSSGVAVVLDGVDPNHFFLSI 191
QY 177 QLRGLEXIAQKKXSNDVAYITTKPEY-KVMN-----DVARGI----- 212
Db 192 TPDGRKKIEMWDFPDBEGFVDTSFSSKVNGLTIINTFYQVAGSLTNESTGLMEYVK 251
QY 213 -----YKADVAGSSGYL-----GQGQIVAVDT 236
Db 252 TYTVSNVITIGNTISANGIYHFGLLPERRYFDLNFQDQDEFFVVLVNSTGNYDIAYVDT 311
QY 237 GLDTRNDS-----SMHEAFRGKTALYALGRTNNAN-----DTNGHGHVAG 279
Db 312 DLDVDFTEBVLPGQYNTVYDVAVFSGYGLPLVYLAELIDPNEBVAVFQDGHGHVAG 371
QY 280 SYLNGXKTN-----KGMAPQANLVFQSIIMDSXG 307
Db 372 TVAGYDSNNDAMDMLSMYSGEWEVFSRLYGMDYTNVTTDTVGVAFGAQMALRVLRS-D 430
QY 308 GLGGLPSNLQTLFQSAKSAGARIHNSGAAANGAVT--TDSRN--YVDYVAKNDMTILF 363
Db 431 GRGSMMDIIEGN-TYAATHGADVISMISLQG--NAPYLDGTDSEVAVDBELTEKYGVVFI 487
QY 364 AAGNEKPNGGTISAPGTAKNATITVGAETELRSPFSYAD-----NIN 405
Db 488 AAGNEKPGNTVGSVGAFTKAITVGAA--VPINGVVYSQALGYDYGYFFYPATNV- 545
QY 406 HVAQSSRGPTKDGRIKPDVMAFGTILSARSSILAPDSFVANHDSKYAVMGCTSMATP 465
Db 546 RIAFSSSGPRIDGEEKPVAVPAGIYSSILPMWIGADF-----MSGTSMATPH 595
QY 466 VAGVAGQLEHNVKRGITTPPSLKALALIGA-----DXGLGYPNGNQWGVVTLDK 519
Db 596 VSGVALLISG-PKEGIIYNPDITIKVLESQATVLRSDPYTGQKRTLELDQHGVLNVTK 654
QY 520 SLNVAVYVNESSKLSTSQKATYXFTATACKPLKISLWSDAPASTTA---SVTLVNDLVLV 576
Db 655 SMEI-----LKAINGTITLPIVDHMDKSVSDPAEVLGVVINGVLAR 696
QY 577 ITAPN-----GTYVGN-----DFKXPKXKNW-----DG-----RNVE----- 605
Db 697 NSIPDIVEMHIKYVDTEYRTFEIATPWIKEPVGSGVILLENNTBFLRVKXVDVEGLEP 756
QY 606 -----NVFINXPGSTYITIEVQAVNVVPGQKFS 634

Db 757 GLYVGRITIDPPT--TPVIEDILNTIVIPKFT 788

RESULT 6

A1341

microbial serine proteinase (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #ext_change 09-Jul-2004

C/Accession: B41341; B41341; S39700; D69730

R/Sloma, A.; Ruffo Jr., G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Pero, J.

J. Bacteriol. 173, 6889-6895, 1991

A/Title: Cloning and characterization of the gene for an additional extracellular serine

A/Reference number: A41341; MUID:92041574; PMID:1938892

A/Accession: A41341

A/Molecule type: DNA

A/Residues: 1-806 <SLO>

A/Cross-references: UNIPROT:P29141; UNIPARC:UPI0000060C20; GB:M76590; NID:g143819; PIDN:

A/Accession: B41341

A/Molecule type: protein

A/Residues: 161-195 <SL2>

A/Cross-references: UNIPARC:UPI0000178D67

R/Glasser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzalez, W.; Hullo, M.F.; Ionescu,

, A.; Rapoport, G.; Danchin, A.

Mol. Microbiol. 10, 371-384, 1993

A/Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr

A/Reference number: JS39655; MUID:95020537; PMID:7934828

A/Accession: S39700

A/Molecule type: DNA

A/Residues: 1-806 <GLA>

A/Cross-references: UNIPARC:UPI0000060C20; EMBL:X73124; NID:g413923; PIDN:CAA51601.1; PI

R/Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.

Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

koeth, J.; Harwood, C.R.; Hensel, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.

Roetter, P.; Koningsreijn, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A.; Laubert, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuho, S.; Maueel

Y. M.; Ogawa, K.; Ogawara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelc

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

A.; Schleich, S.; Schroefer, R.; Scifone, F.; Sekiguchi, J.; Sekowaka, A.; Serot

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpeira, P.; Tognoni, K.; Toseo, V.; Uchiyama,

T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A.; Authore: Yoshikawa, H.F.; Zumbaste, E.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: JS5850; MUID:98044033; PMID:9384377

A/Accession: D69730

A/Structure: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-806 <KUN>

A/Cross-references: UNIPARC:UPI0000060C20; GB:I299123; GB:AL009126; NID:g2636240; PIDN:CA

A/Experimental source: strain 168

C/Comment: The amino terminal sequence of the mature protein and a molecular weight of c

C/Genetics:

A/Genes: vpr

A/Start codon: TTG

C/Superfamily: microbial serine proteinase vpr; subtilisin homology

C/Keywords: hydrolase; serine proteinase

F/1-28/Domain: signal sequence #status predicted <SIG>

F/29-160/Domain: propeptide #status predicted <PRO>

F/180-548/Domain: subtilisin homology #status predicted <SBT>

Query Match 9.3%; Score 286; DB 2; Length 806;

Best local similarity 22.5%; Pred. No. 2-6-12;

Matches 132; Conservative 60; Mismatches 188; Indels 206; Gaps 19;

Qy 155 TKGASXVLAVALDTKXNKEVQ--LRGI-----EXIAQXXXSNDVXYITAKPEYKVMN 206

Db 87 TKAKKAIK--AVKKGKVNREYGVSGFSMKL.PANBEIPKLAVDVAVVPNTYTKDN 144

Qy 207 DVAGIYADVAQS-----SYGLYGQGIYAVADTGLDTER----- 242

Db 145 MKDKVITISEDAVSPQMDASAPYIGANDAMDGLGYTGKIGKIAIIDTGVENHPDLKKNFG 204

Qy 243 -----NDSSMHEARRGKITALYALGRNNANDTNGHTRVAGSILGNGXTYKGMAP 293

Db 205 QYKGYDVPVNDNDYDPKREPTLG-----DPRGATDHGTVAAGVAA--TIKGVAP 252

Qy 294 QANLVFQSIQMSXGGLGGLPSNLQTLFQAXSAGARITHTNSMGAIVNAGVATYTTDSRVNDY 353

Db 253 DATLLAVRDLRG--SSTENNVAGVRAVQDGDVWNLISGNSLNPDMATSTAL-DW 309

Qy 354 VRKNDMTLFAAGNEXPNGTISAPGAKNAITVGATE---NLRPSFGSY----- 400

Db 310 AMSEGVAVVTNSGNSGPMGTGSPGTSREASIVGATQLPLNEVAVTFGSSYSSAKVMGYN 369

Qy 401 -----ADNI----- 404

Db 370 KEDVYKALNKEVELVEAGIGBANDFECKDLTGKAVVVKSGSLAFVDRADNAKAGALGM 429

Qy 405 -----NHWAP 410

Db 430 VVYNNLSGEIEANVPKMSVPTIKSLBDEGKLVSLKAGETKTPKLVSKALGEQVADF 489

Qy 411 SSRGPTKD-GRIKPDVMAPTXILSARSLAPDSSFWANHDSKYAVWGTSMAPTVAGN 469

Db 490 SSRGPTVMDTMKIDISAPGVNIVSTIPTHPD-----HFYGYGSKQGTSMASPHIGA 543

Qy 470 VAOLREHVKNRGITPKRSL--LKAAIAGAA---DXGLGYPMNGOGKRVTLIDKSLNV 523

Db 544 VAVTIQ-----AKPKASVEQIKALNMTVTLTKDSDEGVYFPHNAQAGSARI--RNA 593

Qy 524 AYVNESSXLTSGOKATYXFTATAGKPLKISLWSDPASTASVTL 569

Db 594 --IKRADSLVSPGSGSYGVFLKENGNETKNERFTIENGSIKSYTL 637

RESULT 7

S25835

subtilisin (EC 3.4.21.62) precursor - Bacillus sp. (strain T41)

C/Species: Bacillus sp.

C/Date: 22-Nov-1993 #sequence revision 20-Feb-1995 #ext_change 05-Oct-2004

C/Accession: S25835

R/Davall, S.; Feller, G.; Narinx, B.; Gerday, C.

Gene 119, 143-144, 1992

A/Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph Bacillu

A/Reference number: S25835; MUID:99012966; PMID:1398082

A/Accession: S25835

A/Molecule type: DNA

A/Residues: 1-419 <DAV>

A/Cross-references: UNIPARC:UPI000008B77A; EMBL:X63533; NID:g40198; PIDN:CAA45096.1; PID

C/Superfamily: Subtilisin; subtilisin homology

C/Keywords: extracellular protein; hydrolase; serine proteinase

F/1-23/Domain: signal sequence #status predicted <SIG>

F/24-110/Domain: propeptide #status predicted <PRO>

F/111-419/Product: microbial serine proteinase #status predicted <MAT>

F/135-373/Domain: subtilisin homology <SBT>

F/144,184,359/Active site: Asp, His, Ser #status predicted

Query Match 9.1%; Score 279.5; DB 1; Length 419;

Best local similarity 29.6%; Pred. No. 3e-12;

Matches 97; Conservative 47; Mismatches 133; Indels 61; Gaps 16;

Qy 171 QXNKEVOLRGEXIAQXXXSNDVXYITAKPEYKVMNVA-----RGIVKADVQSS 221

Db 81 QNNKLVTEKPEL-----EITATNKP--ALYNNMAASQSPWGIKALYNNSLTST 132

Qy 222 YGLYGQGIYAVADTGLDTERGNSM--EAFRGKITVYALGRT--NNANDTNGHTRV 277

Db 133 SG--GAGINIAVLDTGVATNHPDLSSNVQCKD-----FTVGTFTTNSCTDRGSHGTHV 185

Qy 278 AGSVLGNKXTNK---GMAPOANLVFQSIM--DSYGLIGLPSNLQTLFQAXSAGARITHT 332

Db 186 AGSALANGGTGSGYVAVLADDDSGYADDAEALRHAGDQTTALNTYVI 245

Qy 333 N-SWGAIVNGAYTTDSRVNDYVRKNDMTILFAAGNEXPNGTISAPGAKNAITVGATE 391

Db 246 NMSIGSSGESSLIT---NAVDYAVDKGVLLIAAAGNSGPRGSGTGYGALVMAVAALAE 302
Qy 392 NLRPSGVSADNINNAVQFSSRGPTKDG-----RIKPDVMAPTXTLSARSSLAPDSF 445
Db 303 NTIQN-GTY-----RVADPSSRGHKRYAGDYVIQKGDVEISAPGAAYST----- 346
Qy 446 WANHDSKYAVMGTSMTPTIVAGNVAQL 473
Db 347 W--PDGATITSGTSMASPHAGLAIAKT 372

RESULT 8

G83753
subtilisin-type protease (EC 3.4.21.-) vpr precursor [similarity] - Bacillus halodurans
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_rev1sion 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: G83753
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; PMID:20512582; PMID:11058132
A/Accession: G83753
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-799 <STO>
A/Cross-references: UNIPROT:Q9KEM1; UNIPARC:UPI000000C39DC; GB:AF001510; GB:BA000004; NID
A/Experimental source: strain C-125
C/Genetic8:
A/Genes: vpr
C/Suprafamily: microbial serine proteinase vpr; subtilisin homology
C/Keywords: hydrolase; serine proteinase
P.1-29/Domain: signal sequence #status predicted <SIG>

Query Match 9.0%; Score 276.5; DB 2; Length 799;
Best Local Similarity 23.5%; Pred. No. 1.2e-11;

Matches 149; Conservative 61; Mismatches 198; Indels 227; Gaps 25;

Qy 77 LETVPAANKLHIXQFNGP-ILR-----ETQKLEKYGAKILDYIPDYAVYVEYE 124
Db 50 LETV-----IVEIEDPSIIIEAKHGOKOSKHELKQAROSVIEQIDLVF--SSTVTAE 100
Qy 125 GPVSPKXXIIEHVESVPEYLPYKXIDPOLFTKGASLVAKXALDTPQXNKVQLRIEIKI 184
Db 101 YPLPSGPALE-----LPAHQIPSLIGIDGVAHVYNNIEYEVTPDEY----- 144
Qy 185 AOXKXSNVXYITAKPEYK-VANDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGHN 243
Db 145 -----VIEKQAVPEMLDSAPFICANDAMEAGF--TGEGITVAIILDITGVDTYHP 191
Qy 244 DSSMEAPFGKITALYALGRTNNANDT-----NGHTVAGSVLGNKTNKMAPQ 294
Db 192 D--LVHAF-GDYKGMDFINDNDPOETPPGDRGIEHTTGTHTVAGTVAANGLI-KGVAD 247
Qy 295 ANLVQSIDXSGXGGLPSNLQTL--FSQAYSAGARITHTSWGAHVNAAYTTDSNVVD 352
Db 248 ANLVLRVY-----GREGSGSTAGVAGIERAVODGADIMLSLANTLNDPFASTAL-D 302
Qy 353 YVRKNDMTILFAAGNEXPNGGTISAPGTAKNAITVGTATENLRPSFGSY-----ADNINH 406
Db 303 WAMAGVAVLVTSNGNSGPRNMTVGSPTGRDAISVQAT--RLPYKRYKASVFTDGDIDY 359
Qy 407 ----- 406
Db 360 PSADIMGPSDELELDGETEYEAFAAGIGKPDGEGVDEGKIALIVGRIPFVEKAEN 419
Qy 407 ----- 406
Db 420 AKAAAGVAILNNVAGVQPTVGLAIPITMLSNEDGLKRNLENGONTTFTSIEFDKL 479
Qy 407 -----VAQSSRGPT-KDGRIRPDVMAPTXTLSARSSLAPDSFMANHDSKYAVMGTSM 461
Db 480 VGEIVADPSSRGPRVMTMIKPDVSAFGVAIVSTIPTHQPDPPY-----GIGSGNGTSM 533

Qy 462 ATPYVAGNVAQLREHVVKRGITPPKSLKALIAQA-----DXGLGYENGQGWGRVTL 517
Db 534 ASPEHVAQAALALIEAH-PWNGV-----DHVYDALMTNIEVLVDENGRVYPHNTQAG----- 584
Qy 518 DKSINAVYVNESXLSSTOKATY-XETATAGKPLX 551
Db 585 --SIRIVDAIESETLVTPQSHSPGTYTKRGKQVE 617

RESULT 9

S23407
subtilisin (EC 3.4.21.62) 1 precursor - Bacillus sp. (strain TA39)
C/Species: Bacillus sp.
C/Date: 04-Dec-1992 #sequence_rev1sion 04-Dec-1992 #text_change 05-Oct-2004
C/Accession: S23407
R/Marinx, B.; Davail, S.; Feller, G.; Gerday, C.
Biochim. Biophys. Acta 1131, 111-113, 1992
A/Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarctic
A/Reference number: S23407; PMID:92256481; PMID:1581352
A/Accession: S23407
A/Molecule type: DNA
A/Residues: 1-420 <NAR>
A/Cross-references: UNIPROT:P28842; UNIPARC:UPI0000136183; EMBL:X62369; NID:940200; PID
C/Genetic8:
A/Genes: sub1
C/Suprafamily: Subtilisin; subtilisin homology
C/Keywords: extracellular protein; hydrolase; serine proteinase
P.1-23/Domain: signal sequence #status predicted <SIG>
P.24-111/Domain: propeptide #status predicted <PRO>
P.112-420/Product: microbial serine proteinase #status predicted <MAT>
P.136-374/Domain: subtilisin homology <SB1>
P.145,185,360/Active site: Asp, His, Ser #status predicted

Query Match 9.0%; Score 274.5; DB 1; Length 420;
Best Local Similarity 28.3%; Pred. No. 6.9e-12;

Matches 107; Conservative 56; Mismatches 142; Indels 73; Gaps 21;

Qy 167 LDTKXN--KEVOLRGIEIXIAQXXXSNVXYITAKPE--YKVM---NDVARGIV---KA 215
Db 73 MKBKQFNLKKNKNTLVEKPEL-----EIAATIDPBLNMAAASQTPWGIKAIYNS 128
Qy 216 DVAQSSYGLYGQGIIVAVADTGLDGRNDSMEAPFGKITAL--YALGRT---NNANDT 270
Db 129 SITQTS---GGGGINIAVLDTGVNTN-----HPDLRNNVBOCKQFTVGTYYTNNSCIDR 179
Qy 271 NGHGRHVAGSVLGNKTNK---GMAPOANLVQSIM--DSXGGLGSLPSNLQTLFQAS 325
Db 180 QGHGRHVAGSALADGCTGNGVGVAPDADLVAKVAGDGSYADIDIAAIRHAGDQATA 239
Qy 326 AGARITHTN-SWGAHVNAAYTTDSRVNDVYVRKNDTILFAAGNEXPNGGTISAPGTAKNA 384
Db 240 LNTKVVIMVSSGSSSESLITNAVN---YSNVKGVLLIAAGNSGPRYGSGTIPGALVNA 296
Qy 385 ITVGATENLRPSFGSYADNINNAVQFSSRGPT-KDG-----RIKPDVMAPTXTLSARSS 438
Db 297 VAVVALEN-KVENGTY-----RVADPSSRGVSWTGDVAIQGDVEISAPGAAYST--- 347
Qy 439 LAPDSFMANHDSKYAVMGTSMTPTIVAGNVAQLREHVVKRGITTPPSL-----L 490
Db 348 -----W--PDGATITSGTSMASPHAGLAIAKIAQVPSASNVVDEGLYRAYENDI 398
Qy 491 KAAITAGAAD--XGLGY 505
Db 399 LSGYVAGYGDDPAGSGRF 416

RESULT 10

S11890
serine proteinase (EC 3.4.21.-) precursor, extracellular - Xanthomonas campestris pv. ca
N/Alternate names: subtilisin-related proteinase
C/Species: Xanthomonas campestris pv. campestris
C/Date: 21-Nov-1993 #sequence_rev1sion 07-Feb-1997 #text_change 31-Dec-2004

C:Accession: S11890
R:Lin, Y.N.; Tang, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
Mol. Gen. Genet. 220, 433-440, 1990
A:Title: A multipurpose broad host range cloning vector and its use to characterize an
A:Reference number: S11890; MUID:90251253; PMID:2187155
A:Accession: S11890
A:Molecule type: DNA
A:Residues: 1-580 <LTI>
A:Cross-references: UNIPROT:P23314; UNIPARC:UPI000012A398; EMBL:X51635; NID:g48533; PIDN
A:Experimental source: Xanthomonas campestris pv. campestris
A>Note: The sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-ALA
C:Superfamily: extracellular alkaline serine protease precursor, Alteromonas type
C:Keywords: extracellular protein; hydrolase; serine proteinase
F:1-32/Domain: signal sequence #status predicted <SIG>
F:168-423/Domain: subtilisin homology <SBT>

Query Match 8.9%; Score 273.5; DB 2; Length 580;
Best Local Similarity 22.7%; Pred. No. 1.3e-11;
Matches 159; Conservative 80; Mismatches 250; Indels 211; Gaps 30;

QY 1 MRKKVPLSVLSAAILSTVALKNSAGAKRFDLDFKGIOTTTXXGFSKXQTGAAAF 60
DB 6 LKRRGSLTTLGASALTSLMLMPAFAG----EYLDGLATVQTHQKFTVTVYKDGSTAL 60
QY 61 LVSEENVTLKXGKXKGLKLTVPANNGLHTXQFNGPILERTKOKLXTGAKIIDIYDVAVI 120
DB 61 ASPG---ALTTSLRTAARVPA-----KAGKALGLNSVRRLLAGPELV-- 100
QY 121 VEYEGDVAXXXXXIHVESVEPEYLPXYXIDPOLFTKGA SXLVKAYALDPTKQXKEVLRG 180
DB 101 -----ALTTSLRTAARVPA-----RADRALDRAEFTLMRQLAADPNVGSVEV---- 129
QY 181 IEXTAQQXXXSNDVXYITAKP-EYKVMNDVAGIYVA--DVAQSGLYGGQGIYVAVDTG 237
DB 130 -----DQILATLTLPNDTRLSEQWAFGTNAGLIRPAMDKATGSGTVAVVDTG 179
QY 238 L-----DTGRNDSMHEAFRGKITLALYALGRTNNAND-----TNGHGT 275
DB 180 ITSHADLANIAGIDFLSDATTAADGNGRDSNADDEGDVAAECGACIPAASSWHT 239
QY 276 HVAGSV--LGNGXTN-KGMAPQANLVFQSIINDSXG-----LGGIPSNL 316
DB 240 HVAGTVAALVNTTGAAGTAAAGAKVAVRVLGKCGSLSDIADALVMASSGTVSGIPANA 299
QY 317 Q--TLFSQASAGARIHTNSMGAIVNGAYTTDSRVNDVYRKNDTILPAANEKPNCGT 374
DB 300 NPAEYINNSLGGGSCSTTMON-AINGAVSRGT-----TVVVAAGNDASVNSG 346
QY 375 ISAPGTAKAATVGAATEN--LSPFSGYADNINHVAQFSSRGPTKGRHKPDVMAPGTXI 432
DB 347 -SLPANCANVAVVAATTSAGAKAATSNFTGT-----DVSAFGSI 386
QY 433 LSARSS--LAPDSFMANHDSKYAYMGTSMTPIVAGNVAQLRHFYKNGRGTTPK--PS 488
DB 387 LSTINSGTTTPOS-----ASYASNGTSMASPHVAGVALVQS--VAPRLTAAVET 437
QY 489 LLK--AALIAAAGXGLG--PENGQGRVTLDSLVNAVYVNSSX 531
DB 438 LLKNTARALPGAGSGGAGIVNADAATVAAINGSGGGGGGNTLTNGPTVGTG---- 493
QY 532 LSTSQKATYFTATAGKPLKISLVMSDAPASTTASVTL--VNDLDVY--TAPNGTX 585
DB 494 AATGELAVYITVPAAG-----SGTLTVTSSGSGGDLVVRAGSAPTDAY 539
QY 586 VGNDFKXKXKXNWDGRNVENVFINKPQSGYTIIEVQAVN 625
DB 540 TCRPYRS-----GNAETCTITAP-SGYTVVLTAKYS 569

RESULT 11
JC4908
alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.
N/Alternate names: subtilase

C:Species: Alteromonas sp.
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 05-Oct-2004
C:Accession: JC4908
R:Tanjido, H.; Miyamoto, K.; Tanaka, K.; Kaidzu, Y.; Inada, C.; Okami, Y.; Inamori, Y.
Biosci. Biotechnol. Biochem. 60, 1284-1288, 1996
A:Title: Cloning and sequence analysis of a protease-encoding gene from the marine bacte
A:Reference number: JC4908; MUID:97141200; PMID:8987544
A:Accession: JC4908
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-715 <TSU>
A:Cross-references: UNIPROT:P70765; UNIPARC:UPI000017A928; DBJ:D38600; NID:g1536787; PI
A:Experimental source: Xanthomonas campestris pv. campestris
C:Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-sensi
C:Genetics:
A:Gene: aprI
C:Superfamily: extracellular alkaline serine protease precursor, Alteromonas type
C:Keywords: hydrolase
F:1-40/Domain: signal sequence #status predicted <SIG>
F:151-456/Product: alkaline serine protease I #status predicted <MAY>
F:182-452/Domain: subtilisin homology <SBT>
F:497-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 8.8%; Score 268.5; DB 2; Length 715;
Best Local Similarity 25.8%; Pred. No. 3.8e-11;
Matches 142; Conservative 51; Mismatches 201; Indels 157; Gaps 28;

QY 168 DTKQXKEVQLRGLEXIXKQXXSNDVXYIT-----AKREYKMD-----VARGIV 213
DB 117 DKLSAKGAQAFMEVVA-----SGNVEIETIDQMLKPPATNDRYNDHYTEAAAGI- 171
QY 214 KADVAQSSYGLYGQGIYVAVDTG---LDTGRNDSMHEAFRGKITLALYALGRTNNAND 269
DB 172 NAPAAMK--ATGGGVVAVVDTGVRPHLDANILPGYDMSITFVANDGARDNDARD 229
QY 270 -----TNG-----HGHVAG--SYGNGXITKMGAPQANLVFQS 301
DB 230 PGDAVTRGECGTSSGQVPVPAQDSSWHGTHVAGTAAVTVNGBGAVGVADKAVFVR 289
QY 302 IMDSXGIGLGPNSLQTLFSQASAGARIHTNSMGAIV-----NGAYTTDSRVNDVY 354
DB 290 VL-----GKCGGLTSDIADIIWASGSDRVPANANPAVYINNSLGGGACATTTQNALNQA 346
QY 355 RKNDTILPAANEKPNCGTISAPGTAKAATVGAAT--ENIRPFGSYADNINHVAQFSS 412
DB 347 RNNGTVIYIAAGNDNDSNANTY-PGNCNGVNVVASVGDGSRAYSYNGANI----- 397
QY 413 RGPTRKDGIRKPDVMAAPGTXILSARSLAPDSFMANHDS-----KYAYMGTSMATP 464
DB 398 -----DVAAPG--GAQSPADDEPGILSTHNSGSGAPSNDSYHSQGTSMAP 442
QY 465 IVAGNVAQLR-----EHFVKN--RGITKPSILKALTI--AGAADXGLG---Y 505
DB 443 HVAGVAAALIKQAKSAPRDEVTILKTTTRSFPASSCNGCGVDAADAAVVAALGDVTP 502
QY 506 PNEGQGRVTLDSLVNAVYVNSSXKLSQKATYFTATAGKPLKISLVMSDAPASTTA 565
DB 503 PTGN-----TLRD--GVAKTGLSGAAGSNQ--FFTEVPAGK-----TNV 538
QY 566 SVTL--VNDLDVITAPNGXVYVNDPKYKXKXNWDGRN-----NNVENVFINKPQSGY 618
DB 539 TPTMSGGTGADLDVYK-----LGSQ--PTSSSDCPRYEGGAEVCSFPAQAGTYH 588
QY 619 IEVQAVNPVG 629
DB 589 VMINGYKAYSG 599

RESULT 12
F69730
cell wall-associated protein precursor wprA - Bacillus subtilis

C:Species: *Bacillus subtilis*
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: F69730
 R:Kunitz, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bercet
 C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, F.; Carter, N.M.; Chd
 A.; Ehrlich, S.D.; Emerson, P.T.; Eutlian, K.D.; Eyring, J.; Fabre, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Poulter, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
 tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Lardinois,
 A.; Luthors, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maseel
 Y, M.; Ogawa, K.; Ogihara, A.; Oudea, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivaletta, C.; Rocha, B.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
 A.; Schuch, M.; Tanaka, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tostato, V.; Uchiyama,
 T.; Winters, P.; Wpat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumelein, B.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: F69730
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-894 <KUN>
 A:Cross-references: UNIPROT:P54423; UNIPARC:UP10000030283; GB:299109; GB:AL009126; NID:9
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: wprA
 F:453-664/Domain: subtilisin homology <SBR>
 Query Match 8.6%; Score 262; DB 2; Length 894;
 Best Local Similarity 23.9%; Pred. No. 1.5e-10;
 Matches 134; Conservative 92; Mismatches 211; Indels 124; Gaps 28;
 49 FSKQQTGAAP--LVSENVTLKXGLKQRTVPA-----NNKLHXYFN---G 93
 245 FDDVSENGASSSYKTEKQKAIINLYDKALQSVSFLKEIKQADRNMQLQKTKAG 304
 94 PILER---TKQKLETKAKILIDYIPDYIVY---ETEG-----DYSKXXX 133
 305 ALITNNNTAAKSEVQIT--KVI PKYDKNSLSVHNEMKGFSAOSKDKISNVKAKKL 362
 134 IEHVESVEPYLPXYKIDPOLPTKGASXLKAXALDTKONKEVQLRGIEIXAO--XXXSND 192
 303 PDNLVSPF--LPKDKQKCAITYASAKRYKSAALITISKSNVFP---AEVDEYKSLAND 416
 193 VXYITAKPEYKVNNDVARGIVK--ADVA--QSSYGLYQCGQ---IYAVADTGLDTRNDSS 246
 417 IQY---PYQWPLKNNNGENGKGVNADYKVPANTLLSKKRLANDTLAVVDYGVDSLAD-- 471
 247 MHBAKRGKITALYA--LGRTNANDTNGHGTIVAGSVLG--NGXTKGAAPQANLVFQ 300
 472 ---LKGKVRITDLGHNPFVGRNNNAMDQGHGTIVAGIIAOSDNGSMTGLNAKAKIIPV 527
 301 SIMDSXGGLGLPSNIQTLFSGQASGAGRIHTNSGAVNGAYTTDSRNVDDYVAKNDMT 360
 538 KYLDSAG--SGPTBQIALGICIKYADKAKAVINLSG---GGISVLEPALKTAADKXVL 581
 361 ILPAAGNEXPNNGTISAPGTAKNAITVAGTENTRPSFSYADNINHVAAQSSRGPTKGR 420
 582 IAAASGNDGEN--ALSYPASSKYVMSVGT-----NRMDMTDFSYVYGL--- 625
 421 IKPDVMAQGYTILASRSLAPDSSFWANHDSKRAYMGTSMAITPIVAGNVAOLREHFVYN 480
 626 ---DIPAGSDI---PSLVPVGN-----VVTMSGTSMTPYAAAAAGILFPAQNPKL 670
 481 RGIPTKPSILKRA--LIAGADXXG-----LGPV-----NGNGMGKRVTLTKSLN 522
 671 KRTKVBMLKTKTADISFESVDGGEELVYDIDPIETPKTPGVDMHSGYGLNWKAVS 730
 523 VA---YVNESSKSLTSQKAT 539
 731 AADLQKVNKLESQTAVRGS 751

RESULT 13
 T29090
 Surface layer-associated STABLE protease - *Staphylococcus marinus*
 N:Alternate names: hyperthermostable protease
 C:Species: *Staphylococcus marinus*
 C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
 C:Accession: T29090
 R:Mayr, J.; Lupas, A.; Kellermann, J.; Eckerskorn, C.; Baumeister, W.; Peters, J.
 Curr. Biol. 6, 739-749, 1996
 A:Title: A hyperthermostable protease of the subtilisin family bound to the surface layer
 A:Reference number: Z2059; MUID:96585442; PMID:8793300
 A:Accession: T29090
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1345 <MAY>
 A:Cross-references: UNIPROT:Q54437; UNIPARC:UP10000062278; EMBL:U57968; NID:91374755; PI
 A:Experimental source: strain F1
 C:Function:
 A:Description: probably serves an exodigester function related to the organism's energy
 A:Note: sticholomeric S-layer component
 Query Match 8.5%; Score 260.5; DB 2; Length 1345;
 Best Local Similarity 29.1%; Pred. No. 3.3e-10;
 Matches 98; Conservative 41; Mismatches 105; Indels 93; Gaps 14;
 251 FRGKITALVALGRTNNANDTNGHGTIVA-----GSVL---GNGXTNK--GMADQANLV 298
 445 YQGRYAL-----VDFPHGHTSVATVYASRRLVYDLYGDDKLYRIMGAPQAKI- 495
 299 FQSTDSXGGLGLPSNIQTLFSGQASG-----ARHTNSW 335
 496 -----AGDAMLLGNITLV--EAMLAGFNIIVTEEBDGYVYLSIDPFGPHRADITISNW 546
 336 GAAVNGAYTTDSRND-----DYVRKNDMTILPAAGNEXPNNGTISAPGTAK 382
 547 GSTIYNFNLQOPFGIDYRSSFMDELAIKRNLYIGDHVITVPAAGNEXPNNGTISAPGTAK 606
 383 NAITVGAATE--NLRSFG--SYADNINHVAAQSSRGPTKGRIRKPDVMAFP-----TX 431
 607 LVITGASTLMDYRIRIYVPEGVAD--EVIAPFSRGRTGGVPRPDVNIAGAFEMASTR 663
 432 ILSARSLAPDSSFWANHDSKRAYMGTSMAITPIVAGNVAOLREHFVONRGITTPPSLIK 491
 664 TIDRGYGAQPDVF-----GTSSEATPTYSGLTALVFOAYKEVYVTTDPDVTAK 712
 492 AALTAGADXXGGLGYPNNGQWGRVTLTKSLNVAVYNE 528
 713 IILKSSAKD--IWIPIAFSGSGRDALKKADVTVFISB 747
 RESULT 14
 A33973
 high-alkaline serine protease (EC 3.4.21.-) yab precursor - *Bacillus* sp. (strain yab)
 N:Alternate names: alkaline elastase; subtilisin yab
 C:Species: *Bacillus* sp.
 C:Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 05-Oct-2004
 C:Accession: A33973
 R:Kaneko, R.; Koyama, N.; Tai, Y.C.; Huang, R.Y.; Yoda, K.; Yamaoka, M.
 J. Bacteriol. 171, 5232-5236, 1989
 A:Title: Molecular cloning of the structural gene for alkaline elastase yab, a new subti
 A:Reference number: A33973; MUID:89359181; PMID:2670913
 A:Accession: A33973
 A:Molecule type: DNA
 A:Residues: 1-378 <KAN>
 A:Cross-references: UNIPARC:UP10000129E78; GB:M88537; NID:9341960; PID:AAA87324.1; PID:
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete
 A:Accession: B33973
 A:Molecule type: protein
 A:Residues: 11-164; 326-355 <KA2>
 A:Cross-references: UNIPARC:UP10000175C84; UNIPARC:UP10000175C85
 C:Genetics:
 A:Gene: ale

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rnp) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

QY 161 IEXIAOXXXXNDVXYITTAKEPEKVMNDVARGIVKADVAGSSYGLYGQGOIYAVADTGIDT 240
Db 161 IEBIAQVYASNDVHHITTAKEPEKVMNDVARGIVKADVAGSSYGLYGQGOIYAVADTGIDT 240
QY 241 GRNDSMHEAFRGKITTAALYALGRTNANDTNGHGHVAGSVLGNKXTKGMAPQANLVFQ 300
Db 241 GRNDSMHEAFRGKITTAALYALGRTNANDTNGHGHVAGSVLGNKXTKGMAPQANLVFQ 300
QY 301 61MDSXGGLGGLPSNLQTLFSQASAGARIHTNSGGAANGAYTTDSRNVDDYRKNDMT 360
Db 301 61MDSXGGLGGLPSNLQTLFSQASAGARIHTNSGGAANGAYTTDSRNVDDYRKNDMT 360
QY 361 ILFAAGNEXPNNGGTISAPGTAKNAITVGAATENLRPSFSYADNINHVAQFSSRGPTKQGR 420
Db 361 ILFAAGNEXPNNGGTISAPGTAKNAITVGAATENLRPSFSYADNINHVAQFSSRGPTKQGR 420
QY 421 IKPDVMAFGTILSARSSILAPDSSFMANHDSKYAYMGTSNATPIVAGNVAQLREHFYKN 480
Db 421 IKPDVMAFGTILSARSSILAPDSSFMANHDSKYAYMGTSNATPIVAGNVAQLREHFYKN 480
QY 481 RGITPKPSILKRALIAGAADKGLGYPNGOGMGRVTLDKSLNAVAYNESSLSTSQKATY 540
Db 481 RGITPKPSILKRALIAGAADKGLGYPNGOGMGRVTLDKSLNAVAYNESSLSTSQKATY 540
QY 541 XFTATAGKPLKISLWSDAPASTTASVTLVNDLDTITAPNGTXYVGNDFXXKXNDG 600
Db 541 XFTATAGKPLKISLWSDAPASTTASVTLVNDLDTITAPNGTXYVGNDFXXKXNDG 600
QY 601 RNVENVFINXPOSGTYTIEVOAYNVVPVGPQFSLAIYN 639
Db 601 RNVENVFINXPOSGTYTIEVOAYNVVPVGPQFSLAIYN 639

RESULT 2

Q93UV9_9BACI PRELIMINARY; PRT; 640 AA.
AC Q93UV9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Protease.
GN Name=PROF;
OS Bacillus sp. KSM-KP43.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=109322;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KP43;
RA Itoh S., Saeki K.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB051423; BAB55674.2; -; Genomic_DNA.
DR PDB; 1WMD; X-ray; A=207-640.
DR PDB; 1WME; X-ray; A=207-640.
DR PDB; 1WMF; X-ray; A=207-640.
DR MEROPS; S08.123; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:protease activity; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILAS_HIS; 1.
DR PROSITE; PS00138; SUBTILAS_SER; 1.
DR HydroLase; Protease; Serine protease.
SQ SEQUENCE 640 AA; 67991 MW; 48BAF7E9D592C15 CRC64;

Query Match 98.8%; Score 3022; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 9.6e-210;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 2 RKKKVFILSVSAAAILSTVALXNPSAGARFEDLPFGIOTTTPDXGSKOXOTGAAL 61
Db 3 KKKCVFLSVLSAAAILSTVALSNPSAGARFEDLPFGIOTTTPAKGSKOGTGAAL 62
QY 62 VESENVKXLXKGLXKXLETPVANNLHLIXQFNGLIIEETKQXLEXTGAKLIDYIPDYAYIV 121
Db 63 VESENVKLPKGLQKLETPVANNLHLIIQFNGLIIEETKQXLEXTGAKLIDYIPDYAYIV 122
QY 122 EYEGDVASXXXXIHHVSEVEPLPYXIIDPOLFTKGASXLKAYALDTKXNKEVOLRGI 181
Db 123 EYEGDVASATSTIEHHVSEVEPLPYXIIDPOLFTKGASELVKAYALDTKQXNKEVOLRGI 182
QY 182 EIXIAOXXXXNDVXYITTAKEPEKVMNDVARGIVKADVAGSSYGLYGQGOIYAVADTGIDT 241
Db 183 EIQIQFPAISNDVLYITTAKEPEKVMNDVARGIVKADVAGSSYGLYGQGOIYAVADTGIDT 242
QY 242 RNDSMHEAFRGKITTAALYALGRTNANDTNGHGHVAGSVLGNKXTKGMAPQANLVFQ 301
Db 243 RNDSMHEAFRGKITTAALYALGRTNANDTNGHGHVAGSVLGNKXTKGMAPQANLVFQ 302
QY 302 1MDSXGGLGGLPSNLQTLFSQASAGARIHTNSGGAANGAYTTDSRNVDDYRKNDMTI 361
Db 303 1MDSXGGLGGLPSNLQTLFSQASAGARIHTNSGGAANGAYTTDSRNVDDYRKNDMTI 362
QY 362 LFAAGNEXPNNGGTISAPGTAKNAITVGAATENLRPSFSYADNINHVAQFSSRGPTKQGR 421
Db 363 LFAAGNEXPNNGGTISAPGTAKNAITVGAATENLRPSFSYADNINHVAQFSSRGPTKQGR 422
QY 422 KPDVMAFGTILSARSSILAPDSSFMANHDSKYAYMGTSNATPIVAGNVAQLREHFYKNR 481
Db 423 KPDVMAFGTILSARSSILAPDSSFMANHDSKYAYMGTSNATPIVAGNVAQLREHFYKNR 482
QY 482 GITPKPSILKRALIAGAADKGLGYPNGOGMGRVTLDKSLNAVAYNESSLSTSQKATY 541
Db 483 GITPKPSILKRALIAGAADKGLGYPNGOGMGRVTLDKSLNAVAYNESSLSTSQKATY 542
QY 542 FTATAGKPLKISLWSDAPASTTASVTLVNDLDTITAPNGTXYVGNDFXXKXNDG 601
Db 543 FTATAGKPLKISLWSDAPASTTASVTLVNDLDTITAPNGTXYVGNDFXXKXNDG 602
QY 602 NNVENVFINXPOSGTYTIEVOAYNVVPVGPQFSLAIYN 639
Db 603 NNVENVFINXPOSGTYTIEVOAYNVVPVGPQFSLAIYN 640

RESULT 3

Q76L84_9BACI PRELIMINARY; PRT; 640 AA.
AC Q76L84;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Protease.
OS Bacillus sp. KSM-9865.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=192495;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KSM-9865;
RA Okuda M., Saeki K., Kobayashi T.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB084155; BAC82522.1; -; Genomic_DNA.
DR HSP; P00782; 1AQN.
DR SMR; Q76L84; 207-640.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:protease activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR002093; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILAS_HIS; 1.


```

RX MEDLINE=2056675; PubMed=1118284; DOI=10.1006/bbrc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046405; BAB21268.1; -; Genomic_DNA.
DR HSSP; Q45670; IDBI.
DR SMR; Q9AQR4; 1-433.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KM Hydrolyase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45576 MW; 98A2DF18FE60DDC CRC64;

Query Match 63.8%; Score 1952.5; DB 2; Length 433;
Best Local Similarity 86.6%; Pred.No.11e-132;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

QY 206 NDVARGIVKADVAQSSYGLYGQGIYVAVDTGLDTGRNDSMHBAFRGKITALYALGRTN 265
DB 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMHBAFRGKITALYALGRTN 60

QY 266 NANTNGHGTIVAGSVLNGKTNKGMAPQANLVFQSIMDSKGLGGLPSNLQTLFSSQAXS 325
DB 61 NANTNGHGTIVAGSVLGN-ALNKGMAPQANLVFQSIMDSKGLGGLPSNLQTLFSSQAMN 119

QY 326 AGARHHTSMGAAVNGAYTTSDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
DB 120 AGARHHTSMGAPVNGAYTANSRQVDEYVRNNDMTILFAAGNEXPNGGTISAPGTAKNAI 179

QY 386 TVGATENLRPSFGSYADNINHVACFSSRGPTKDGRIKPDVMAPTXILSARSSILAPSSSF 445
DB 180 TVGATENLRPSFGSLADNPNHIAQFSSRGATRDGRIKPDVTAPEGFILSARSSILAPSSSF 239

QY 446 WANHSKAYWNGGSMATPIYAGNVAQLREHFVNKRGITPKPSLIKAAALTAGADXGLGY 505
DB 240 WANYSKAYWNGGSMATPIYAGNVAQLREHFVNKRGITPKPSLIKAAALTAGADTVGLGY 299

QY 506 PNGNGGWRVTLDSLVNAVYVNESSXLSQKATYXFTATAGKPKILSVMSDAPASTTA 565
DB 300 PNGGQGWKRVTLDSLVNAVYNEATALTGQKATYSFOQNGKPKILSLVMTDAPGSTTA 359

QY 566 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXXNMDGRNNVENVFINKPQSGTYTIEVQAYN 625
DB 360 SVTLVNDLDLVITAPNGKYVGNDFSYPDNMDGRNNVENVFINAPQSGTYTIEVQAYN 419

QY 626 VPVGPOKPSLAIVN 639
DB 420 VPSPQRPFSIAIVH 433

RESULT 6
Q9AQR4_9BACT
ID Q9AQR4_9BACT PRELIMINARY; PRT; 433 AA.
AC Q9AQR4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Protease (Fragment).
GN Name=PROA;
OS Bacillus sp. D6.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
```

```

OX NCBI_TaxID=127889;
RN [1] NUCLEOTIDE SEQUENCE.
RP STRAIN=D6;
RC MEDLINE=2056675; PubMed=1118284; DOI=10.1006/bbrc.2000.3931;
RX Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046402; BAB21265.1; -; Genomic_DNA.
DR HSSP; Q45670; IDBI.
DR SMR; Q9AQR4; 1-433.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KM Hydrolyase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45636 MW; 52087E0A2516107F CRC64;

Query Match 63.7%; Score 1948.5; DB 2; Length 433;
Best Local Similarity 86.4%; Pred.No.2.2e-132;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

QY 206 NDVARGIVKADVAQSSYGLYGQGIYVAVDTGLDTGRNDSMHBAFRGKITALYALGRTN 265
DB 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMHBAFRGKITALYALGRTN 60

QY 266 NANTNGHGTIVAGSVLNGKTNKGMAPQANLVFQSIMDSKGLGGLPSNLQTLFSSQAXS 325
DB 61 NANTNGHGTIVAGSVLGN-ALNKGMAPQANLVFQSIMDSKGLGGLPSNLQTLFSSQAMN 119

QY 326 AGARHHTSMGAAVNGAYTTSDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
DB 120 AGARHHTSMGAPVNGAYTANSRQVDEYVRNNDMTILFAAGNEXPNGGTISAPGTAKNAI 179

QY 386 TVGATENLRPSFGSYADNINHVACFSSRGPTKDGRIKPDVMAPTXILSARSSILAPSSSF 445
DB 180 TVGATENLRPSFGSLADNPNHIAQFSSRGATRDGRIKPDVTAPEGFILSARSSILAPSSSF 239

QY 446 WANHSKAYWNGGSMATPIYAGNVAQLREHFVNKRGITPKPSLIKAAALTAGADXGLGY 505
DB 240 WANYSKAYWNGGSMATPIYAGNVAQLREHFVNKRGITPKPSLIKAAALTAGADTVGLGY 299

QY 506 PNGNGGWRVTLDSLVNAVYVNESSXLSQKATYXFTATAGKPKILSVMSDAPASTTA 565
DB 300 PNGGQGWKRVTLDSLVNAVYNEATALTGQKATYSFOQNGKPKILSLVMTDAPGSTTA 359

QY 566 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXXNMDGRNNVENVFINKPQSGTYTIEVQAYN 625
DB 360 SVTLVNDLDLVITAPNGKYVGNDFSYPDNMDGRNNVENVFINAPQSGTYTIEVQAYN 419

QY 626 VPVGPOKPSLAIVN 639
DB 420 VPSPQRPFSIAIVH 433

RESULT 7
Q9AQR2_9BACT
ID Q9AQR2_9BACT PRELIMINARY; PRT; 433 AA.
AC Q9AQR2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
```

DE Protease (Fragment).
GN Name-PROC;
OS Bacillus sp. Y.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=133779;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkley;
RX MEDLINE=20568675; PubMed=1118284; DOI=10.1006/brc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
evolutionary relationships";
RL Biochem. Biophys. Res. Commun. 279:313-319 (2000).
DR EMBL; AB046404; BAB21267.1; -; Genomic_DNA.
DR HSSP; Q45670; 1DBI.
DR SMR; Q9AQR2; 1-433.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilysis and peptidolysis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Hydrolase; Protease; Serine protease.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 433 AA; 45588 MW; B81291A803C775AB CRC64;

Query Match 63.5%; Score 1941.5; DB 2; Length 433;
Best Local Similarity 86.2%; Pred. No. 76-132;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

QY 206 NDVAGIYVADVAQSSYGLYGGQGVAVADTGLDGRNDSMHEAPRGKITLALYALGRN 265
DB 1 NDVAGIYVADVAQSSYGLYGGQGVAVADTGLDGRNDSMHEAPRGKITLALYALGRN 60

QY 266 NANTDNGHGTAVAGSVLGNKXNKGMAPQANTVPOSIMDSXGLGGLPSNLQTLFSQAXS 325
DB 61 NASDNGHGTAVAGSVLGN-ALANKGMAPQANTVPOSIMDSXGLGGLPSNLQTLFSQAXN 119

QY 336 AGARIHTNSWGAAVNGAYTTDSRNVDDVYRKDDMTLLFPAAGNEXPGGTISAPGTAKNAI 385
DB 120 AGARIHTNSWGAAPVNGAYTANRQVDEYVRNDDMTVLFPAAGNEXPGGTISAPGTAKNAI 179

QY 386 TVGATENLRPSFGSADNINNHVAQSSRGPTDGRKRPVMAPGTILSABSSLAPDSGF 445
DB 180 TVGATENLRPSFGSADNENHIAQSSRGATRDGRKRPVTAAGTIFILSABSSLAPDSGF 239

QY 446 WANHSKYAYMGGTSMATPIVAGNVAQLREHFVKRGITPKPSLLKAALIAAAXGIGY 505
DB 240 WANHSKYAYMGGTSMATPIVAGNVAQLREHFVKRGITPKPSLLKAALIAAAXGIGY 299

QY 506 PNGOGMGWVTLDKSLNVAAYNVSXSLSTSORATYFTATAKPLKISLVMSDAPASTTA 565
DB 300 PNGOGMGWVTLNKSILNVAAYNVAATATATGAKATYSFGQAKGPKLISLVMDAGSTTA 359

QY 566 STTLVNDLDTLTAPNGTXYVGNDRPKXPKKXKMGDGNNTNVPINXPGSGTTTIRVOAYN 625
DB 360 STTLVNDLDTLTAPNGTXYVGNDRPKXPKKXKMGDGNNTNVPINXPGSGTTTIRVOAYN 419

QY 626 VPGVQXPSLAIYN 639
DB 420 VPGVQXPSLAIYN 433

RESULT 8
QANB18_9MICC
ID QANB18_9MICC PRELIMINARY; PRT; 697 AA.

AC QANB18;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Peptidase S8 and S53, subtilisin, kexin, sedolisin.
GN ORFNames=ArthDRaft_0589;
OS Arthrobacter sp. FB24.
OC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
OC Micrococcales; Micrococaceae; Arthrobacter.
OX NCBI_TaxID=290399;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Isaran S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (PGF-ORNL);
RA Laximer F., Land M.;
RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAH0100025; EAL94539.1; -; Genomic_DNA.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Hydrolase; Protease; Serine protease.
SQ SEQUENCE 697 AA; 73821 MW; B892496C67C0714 CRC64;

Query Match 28.2%; Score 863.5; DB 2; Length 697;
Best Local Similarity 34.2%; Pred. No. 116-53;
Matches 230; Conservative 98; Mismatches 215; Indels 129; Gaps 19;

QY 82 ANNKHI-XQFNGPILEBTKOXLEXTGAKIIDYIPDYIYVEBEDVSKXXIXIBRVSV 140
DB 30 ASBSHIIILQTAEPITTAQRAELAGIDVBMGEYSDNTYLAAPADINRRALPFEVMA 89

QY 141 EPLYEXYXIDPQLFTKGA-SKLVRKXALDTQXKREVQ-----RGIEIXAQXXX 189
DB 90 DVSXVFKLPPLPRASADTGVNRSIADHBPDRRLRRVLLHPGIEAGBELIARVAA 149

QY 190 S-----NDVAYITAKREYKAMNDVAKGIYVADVAQ 219
DB 150 AARVPAVAVATPGKLRITTSVGQLELAIDEIHEIPVRERQLENNVAREIINADVQL 209

QY 220 SSYGLYGGQGVAVADTGLDGRNDSMHEAPRGKITLALYALGRN--NNANDNGHGTIV 277
DB 210 NGTTYRGAEEVAVADTGFDTG-DNANPAPATGTVQTLIALGRAPADADPHGHGTIV 268

QY 278 AGSVLGNKXTN-----KGMAPQANTVPOSIMDSXGLGGLPSNLQTLFSQAXSAGARIH 331
DB 269 AGSVLGRNNSATMGAIIGTAPBALILIQSLDPRGIGGIPVNLNDFQKTYDDGARVH 328

QY 332 TNSKGA-AVNGAYTTDSRNVDDIYAKN-DMITLPAAGNEXPNG-----GTISAPGTA 381
DB 329 TNSWVCPGLNLPYDASSREIDFVNNHDPDVCFAAGNDGVGNSDGTVDNSISGSQNA 388

QY 382 KNAIVGATENLR-----SPFSY-----ADNINHVAQSSRGPTDGRK 422
DB 389 KNCITVGAESSELRKEFTSYGTIPGDPAPAPVAKDKKANPDGAVASSSGPTGRIK 448

QY 423 PDVNAFGTIIISABSSLAP-DSSFWANHDSKAYVNGTSMATPIVAGNVAQLREHFVNOR 481
DB 449 PDVNAFGTIIISABSSLAPMGNTFTSTDPLEFFDSGTSMATPIVAGCAAVLRETLVNG 508


```

Db      8 ITVSTAAVLSVA---AAGATVLTKNAG--AIDT-----NKLSTNVAASMMMAKAT 55
Qy      68 KKKKGLKKLETVRANKKLIHQNGPPILEETKQKLEXTGAKILDYIPDYAVIV- YEGD 126
Db      56 E-----YIVQFKKAVTEKDKALKAQ- FEVFGYLPDDALVVRGSIYS 96
Qy      127 VASXXXKIEHVSVEPYLKYXID-----POLFTKASKLY-----KAXA 166
Db      97 LVTFKTNHGVQAVNVKAGNYKNSNPAPASVPTKQNMVAVLVNTFKSHVEVIAKIKR 156
Qy      167 LPTKQKKEVQ-----LRGIEKI-----AQXXSNDYXYIT 197
Db      157 MOVKVAQVVDGHHITALLPRGLVPAVALTYGHHVQPARVIRSLFAPDEDLADYSAT 216
Qy      198 AKPEYKQNDVARG--TVKDAVAQSSYGLYGOQIYAVADTGLDTRNDSMHEAFRGKI 255
Db      217 AAGDYSDLTGDSGSTLNMFPDAAM-AMGYAKGQQTVMADTGLDSG-NTGAIHQDPAGGV 274
Qy      256 TALYALGRNNA-NQTNHGHTVAGSVLNGKTNKCM-----APQANLYFQSIIMDSKGGIG 310
Db      275 IISGYPGLMSKMSWSDPMHGHTVAGSVMGRTASKGLKLGAYEAMVAEGMWSPMKQVL 334
Qy      311 GLPSNLQTLFSQAKSAGARIHTNSMGAA--VNGAYTTDSRNVDDYVAKN-DMTILFAAGNE 368
Db      335 STPSLTGLDFPKAPADGARIHTNSMGCAITFGAYDNPAVQVDEMSYANPDMLILFAAGNS 394
Qy      369 XP-----NGGTISAPGTAKNAITVGA TENI-----RPSFG 398
Db      395 GADKNKQGRIDNSNMASSPGTAKKNVLTVGASENVTKSGGIQVPSIKRAAKDEWSPSRIYS 454
Qy      399 ST-ANININVAQFSSSGPTKQGRIRKPDVAAFGYIILSARSSLPDSSFFAANDSKAYNG 457
Db      455 STISNGNGLAFSSSGPTTDRKTPDIYAPGTNNVLSVPSQKDSPLMGAAYNKQVWSG 514
Qy      458 GTSMAITPIVAGNVAQLREHFVKNRG--TPKPSILKXALLAGADKGLG----- 505
Db      515 GTSMAITPLAAGAAALAKQVLVEKLGKMP SAALMKATMLHTAVDMTPGPGFBI GAARGOE 574
Qy      506 -----PNGNQGWGRVTLDKSLNVA-----YNSSXSLSTQKATYXTATAGKPLKISLV 556
Db      575 ILTRBPNSDGGRVADVANIAMLGATQFVDRNGVAGQAGAFVSEBTLMAPSLVANLVW 634
Qy      557 SAAAPSTTASVTLVNDLIVITAPNGTXYVNDPFXKPKXXKNDGRNNVENVFINKPQSGT 616
Db      635 TDAPSANAAQALVNDLIVITLPNGQTLSPMDHI-----NNLEMIKSGLPACT 684
Qy      617 YTIIVQAVVVPV---GPOKFSL 635
Db      685 YKLIYKGRVPOGKNGAQAYAL 706

RESULT 11
0747P6_GEOSL PRELIMINARY; PRT; 2030 AA.
ID 0747P6_GEOSL PRELIMINARY; PRT; 2030 AA.
AC 0747P6;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Fibronectin type III domain protein.
GN OrderedLocustNames=GSU3219;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.B., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.F., Wu D., Ward N.L., Beaman M.J., Dodson R.J.,
RA Madupu R., Brinkac L.M., Daugherty S.C., Deboy R.T., Durkin A.S.,
RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,

```

```

RA Davidson T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA Weidman J.F., Khouri H.M., Feldblyum T.V., Ulteback T.R.,
RA Van Aken S.B., Lovley D.R., Fraser C.M.,
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RL environments."
RL Science 302:1967-1969 (2003).
DR EMBL; AB017180; AAR35610.1; -; Genomic_DNA.
DR HSSP; P27693; IAH2.
DR TIGR; GSU3219; -.
DR GO; GO:0004289; P:subtilase activity; IBA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IBA.
DR InterPro; IPR01635; AAPH.
DR InterPro; IPR003961; RN_III.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF07705; CARD8; 8.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
KW Complete proteome.
SQ
SEQUENCE 2030 AA; 207213 MW; ED7ADC27DD141B01 CRC64;

Query Match 20.9%; Score 638; DB 2; Length 2030;
Best Local Similarity 29.4%; Pred. No. 8.5e-37;
Matches 188; Conservative 93; Mismatches 243; Indels 116; Gaps 19;

Qy      82 ANKKLIHQFNGPPILEETKQKLEXTGAKILDYIPDYAVIVVEYGDVXSXXXIEHVSVE 141
Db      54 AIRKMWLVQFNGFVAPBQROLEALGCRIGDYMPTNFAVALMDDBAAKRAVALLSPFEDIT 113
Qy      142 PLYLKYXIDPOLFTGASKLYKAXLD-TKQKNGVQLRG----- 181
Db      114 RPAF-----ADKLVTARKDLTAAPTSIRIRIKVLAVDDPADRAAVIAATLKG 161
Qy      182 -----EXIAQXXXSNDVXYITAKPEYKQNDVARGIVKDAVQSG--- 220
Db      162 NGRILNAGARTTVAVPEBELAPLAQOSETAWIGVGBELRLNSDAAVYQTNBYDNRTI 221
Qy      221 -SYGLYGOQIYAVADTGLD-----TGRNDSMHEAFRGKITALLYALGRTNNAVD 270
Db      222 WEKGITGAGQIVGIADSGVDYDMPFADPENGALPGRGRIKYG-----YDATTGDNHDA 276
Qy      271 NGHGTIVAGSVLGN--GXTKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFSQAKSAG 327
Db      277 DGHGTHIGTICGDRGPGKNGIAPGARIVHQDVTGDTGLTG-SLELETYLKXAYDSG 335
Qy      328 ARIHNSMGAAVNGAYTTDSRNVDDYV-RKQNDMTILFAAGNEXPNKGTISAPGTAKNAT 386
Db      336 ARIFGSWGVD-SGNYDALAALDDPSWRHKDFLAVFANGNGPAPBQRTATSPALNNATS 394
Qy      387 VGATENLRPSFGSYADNINHVAAQFSSRGPTKQGRIRKPDVMAFGYIILSARSSLPDSSFW 446
Db      395 VVATGN-----CTDAAT---VSABSSVQAGADGRANPVGAPGQGVNARS---DGLLG 442
Qy      447 ANHDSKAYVMGTSMAITPIVAGNVAQLREHF-----VKNRGITPKESLKAALIA 496
Db      443 SGNSTGMA-MSGTSSVAAAVTSGAALIRQYFDGFFPGSPAPATKRLPSSALLLAVALVN 501
Qy      497 GAA-----DXGLGYNGNQGWGRVTLDKSL-----NVAYVNSSXSLSTQKATYXFLTA 544
Db      502 SBEALLSDPDGDCSPSKGSGMRPCLINTLPNGDSHSLBVVDGCTGLSTGWCQKRLYS 561
Qy      545 TAGKPLKISLVSDAPASTASVTLVNDLIVITAPNGTXYVGNDF-----XXPKXXNM 598
Db      562 PGRRLKITLTAMTDAAPARGATSPLTNDNLVVAAPDQTTLYLGNDLNSHSDYBSRTGTF 621
Qy      599 DGRNNV-ENVFINAPQSGTITIEVQAVNVVPGQXFSIAI 637
Db      622 SDRVAVBEQVVIKRPVAGTYLVKVGASIPVGPPQPFALVM 661

RESULT 12
Q4H0T5_GIBZS

```

ID Q4HUY5 GIBZE PRELIMINARY; PRT; 1088 AA.
AC Q4HUY5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=F011223.1;
OS Glibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocrematiales; Hypocreales; Nectriaceae; Glibberella.
OK NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbbaum C., Abouelleil A., Allen N., Anderson S.,
RA Archetti H.M., Barina N., Bastien V., Bloom T., Boguslavsky L.,
RA Boudgaltier B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal A., Karatas A.,
RA Kellis C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lai A.,
RA Ma L.-Y., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mancini E., McCarthy M., Meldrum J., Meneus L.,
RA Mitrova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Nordu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupia A., Ramsamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schuer S., Schuback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer S., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkatarman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody B.,
RA Lander E.;
RL "Fusarium graminearum genome sequence."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AACM01000460; EAAV5433.1; -!- Genomic DNA.
CC KW Hypothetical protein.
SQ SEQUENCE 1088 AA; 119629 MW; E3B3CB94C07F542 CRC64;
Query Match 19.4%; Score 594; DB 2; Length 1088;
Best Local Similarity 31.8%; Pred. No. 5,7e-34;
Matches 197; Conservative 98; Mismatches 231; Indels 94; Gaps 29;
QY 95 ILLETKKOLEXTGAKILDIPIYAVIVR-YEGDVAXSXXXIIEHVSEVPYLPXYIDPOL 153
DB 476 ILIQKQ--LTSPYLRLKSTISLKMWDRELRTRYRVDCLIEHPNVTAKLAERI 532
QY 154 FTGASXLVKAALDTQKXNKVEQRLGIEKIAQXXXNDVYITAKPEKVMNDVARGIV 213
DB 533 AEKGVDTLK-LAVSPVRIPLTVHDDKLEALKLDSIRIEEV--RPD-EVLNDLARETL 588
QY 214 KADVAQSSYGLYGQCIYAVADTGLDTRNDSSM---HEAFRKITLALYALGRTNAND 269
DB 589 NANITLALSTSEYGNQKVCVADTGFDDKMDDEMILVHAFNGVHEHLALM-LGDSKD 647
QY 270 TNGGTHYAGSVLNGXNTN-----KGMAPQNLVFOSI-----MWSXGLGLSPENT--QT 318
DB 648 TAGGTHVCAISICGGLYKNGDIRVAGAPGATLWVQSIQVSRPNKGAIEVPMDLQ 707
QY 319 LFSQXASAGARIHTNSMGA---AVNG--AYTTDSRNVDDYV-RKNDWTLFPAAGNEXENG 372
DB 708 LFSNPKYKGIHSHSMKWDKATKGQGLGEGQANDIDKFIYDHDPPVVLVAAGNAKA 767
QY 373 GT---ISAPGTAKAATVIGATENLRPSFGSYADN-----INHAQPSSSRGPTD- 418
DB 768 KSKSNHIGAAGSAFNCITVAGTTRPNNDYGFNDNEGAKPMWTRINDTAKFSSRSRPTKFG 827

QY 419 -----GRIKPDVMAPTXILSARS-SLAPDS-----SFWANHDSKYAMGSTMATP 464
DB 828 RDINGNEAGRIKIKDVAAPGAILLSASRAAKOSRRKRVWYGRGDDWMTSGTSMSTP 887
QY 465 IVAGNVAQLREHFYKNGRITPKPSLLKALATAGAD-----XGLGTPNGQNGRVTLD 518
DB 888 LVACCVALLREALEHKEKPSAALIKALLVNGAVNFSQGLGIGY-DYDQGFGRVDID 946
QY 519 KSLVV-----AYVNSSXKLSSTSQ-----KATYPTAT-----AGK-PLKSLVMSD 558
DB 947 SSISMVTKLSFVDGKLPEDTQFPVAPLRQVPSERERTSLIPVPGRRNLITVTLAPD 1006
QY 559 APASTASVTLVNDLIVTAPNCTXYVGNDFXPXPKXXNMGRNVNVPINXFGSGTYT 618
DB 1007 KPAQ---SGLMQNDINILVLS-GGAERHGNMGKPE---GYDHTNNVEKTIENVPGETFK 1059
QY 619 IEVQAY-NVPV-GPOXPSLA 636
DB 1060 IVASIMWNIDVKAPTSFAVA 1079
RESULT 13
ID Q54M84 D1CDI PRELIMINARY; PRT; 1741 AA.
AC Q54M84;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE ABC transporter B family protein.
GN Name=tagC; ORFNames=DOB0191192;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
OK NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Suggang B., Bertram M., Song J., Olsen R., Szafranski K., Xu O.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankler A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Rey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kethornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Murry D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loulsged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rablinowitch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Splegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winkler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.A.,
RA Williams J., Dear P.H., Noegel A.A., Barrett B., Kuspa A.;
RL "The genome of the social amoeba Dictyostelium discoideum";
RL Nature 01-01(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAF01000133; EAL64353.1; -!- Genomic DNA.
CC DR GO; GO:0016021; C:integral to membrane; IEA.
CC DR GO; GO:0005524; F:ATP binding; IEA.
CC DR GO; GO:0016887; F:ATPase activity; IEA.
CC DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
CC DR GO; GO:0000166; F:nucleotide binding; IEA.
CC DR GO; GO:0004289; F:subtilase activity; IEA.
CC DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
CC DR GO; GO:0006810; P:transport; IEA.
CC DR InterPro; IPR003593; AAA_ATPase.
CC DR InterPro; IPR011527; ABC_membrane_1.
CC DR InterPro; IPR001140; ABC_TM_transp.
CC DR InterPro; IPR003439; ABC_transp_like.
CC DR InterPro; IPR000209; Pept_S8_553.

FT	COMPBIAS	1353	1357	Poly-Asn.
FT	COMPBIAS	1358	1364	Poly-Asp.
FT	COMPBIAS	1381	1386	Poly-Asn.
FT	COMPBIAS	1707	1729	Poly-Asn.
FT	ACT_SITE	325	325	Charge relay system (By similarity).
FT	ACT_SITE	372	372	Charge relay system (By similarity).
FT	ACT_SITE	637	637	Charge relay system (By similarity).
FT	CARBOHYD	390	390	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	536	536	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	547	547	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	614	614	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	669	669	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	735	735	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	741	741	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	776	776	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	832	832	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	887	887	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	1251	1251	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	1385	1385	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	1386	1386	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	1454	1454	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	1704	1704	N-linked (GlcNAc. . .) (potential).
SEQ	SEQUENCE	1743 AA;	194146 MW; 12DB363B2F729839 CRC64;	
Query Match				
Beet Local Similarity		16.9%;	Score 518.5;	DB 1; Length 1743;
Matches 194;		Conservative 114;	Mismatches 213;	Indels 281; Gaps 33
QY	HIXONGILBERTKXL	---	EXTGAKILDIYIDVAYIV	---EYBDDVAXX 130
DB	115 YIVQPKRIDINBTRQLEKFLIGTDI	VLDEQPYQSHIVYIPHDSFLVMTQSQVLLSS 174		
QY	131 XXXIEHVESVEPY	-----LPXXYIDPOLFTKGASLTVK	---AXALDT--KQXN 173	
DB	175 KEWWSWIEPEPESNKHLYNNEKSI	GLPYIIL	-----LSQSTMSILQREMENTANSLIKSN 230	
QY	174 KEVQLRGI	-----	EXIAQXXXNDVXYITAKPEYK 203	
DB	231 SKVKLTLLINQKLLKSIYVCNDESP	SPSCSLINSEKLVQWISQESGN	---FIERSSEKQ 287	
QY	204 VMNDVARGIV	-----KADVAQSSYGLYGGQGVYAV	VDLTGIDR---NNS	---245
DB	288 TANRLSPKPVVFCTKDTLVNNDVDIP	-----IRGKQILISLADTGIDGSHCFPSDSKYPI 342		
QY	246 ---SMEHAFRGKITALVALGRTNANDT	GHGTTHVAGSVLQ	-----NGXTNKGMARPOA 295	
DB	343 PLNSVNLNHR	-KVYVYITTTSTDDSDKVDGHDGTH	ICGSAAGCFPDBSSVNISSFGLATDA 401	
QY	236 NLVFGSINDSKXGLGGL	---PSNLQTLFSPQASAGARIHTNSGA	---AVNGAYTTDSRN 349	
DB	402 KIAF	---FDLASGSSSLFPPSDLKQLYQ	LYADAGARVHCDMSGVSVEGYSYSDPAS 458	
QY	350 VDDYVRKN	-DMTILPAAGREXNGGTIS	---AGCTAKMILVATENTLR	-----PSRS 399
DB	459 IDDFLFTHPDPILRLAAGN	---NEQYLSLLTQSTAKNIVTGAHQ	TIHENTYLTDPENYIN 515	
QY	400 YADN	---		---404
DB	516 YQSSVDINQELLCDPSRYCNYTTAQC	CLSEASATTGGLASCCTFLRKSVITDAANTOPPLY 575		
QY	405 ---NHVAFSSRGPTVQGRIKPDVMA	PGTYILSARSSLA	-----PDSSFMANHDSKYVM 456	
DB	576 NENNICSPSSKCPHTDGRMKFALVA	PEYITTSARANGAATTTQCCDGS	L	---PRTNALLA-I 633
QY	457 GGTSAATITIVAGNAQLREH	-----PYKRGITPKPBLKALALIGA	-----498	
DB	634 SGTSAATSPFAAATTLIKQYLV	DGYYPGSIYVESNKLOPTGSLKALMINNAQLINGTFQ 693		
QY	499 ---ADXGLGYPNQ	-----QMGVLTLDKSLNAVYVNESS	-----530	
DB	654 LITSSSTIYPSNQVEENRAGASLV	QMGKAIKMSNLHVVNNNSNNNTKTSIDGCIYKFDGI 753		
QY	531	-----XLSQKATYXFP	-----ATAGKDLK	---ISLVNSDAPASTT 564

DB 754 GGLDLRLVYPNQWKESESLSTGNTSYSCFYKDPSSSSSSNGNNIPRVAVLTWMDPPSYAG 813

QY 565 ASVTLVNDDLYI-----TAPGYTYGVGNDYFXXKXNDGRNVENVFINXP 612

DB 814 AKFNVLNDDLMITYRRDNGSTIFPSSNGGSSFLG-----LAFTDITLNVGCIYVHNPT 867

QY 613 OSGTYTIEVOAYNVVPGQXFS 634

DB 868 BEMTRFMVAGTVMGPNFNS 869

RESULT 15

08T9W1 DICDI PRELIMINARY; PRT; 1825 AA.

ID 08T9W1 DICDI PRELIMINARY; PRT; 1825 AA.

AC 08T9W1

DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)

DE Serine protease/ABC transporter Tagd (ABC transporter B family protein).

GN Name:tagd; ORFNames=DD80191427;

OS Dictyostelium discoideum (Stime mold).

OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

OC NCBI_TaxID=44689;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Ax4.

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases. [2]

NUCLEOTIDE SEQUENCE.

RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A., Tugang R., Bertman M., Song J., Olsen R., Szatranski K., Xu Q., Bankier A.T., Lehman R., Hamlin N., Davies R., Gaudet P., Fey P., Plitcher K., Chen G., Saunders D., Sodergren E., Davis P., Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Baon N., Farbrother P., Desany B., Just E., Morio T., Roat R., Churcher C., Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I., Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsey R., Hauser H., James K., Qutles M., Mohan M.B., Saito T., Buchrieser C., Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A., Louisse H., Mungall K., Oliver K., Price C., Quail M.A., Urushihara H., Hernandez J., Rabinowitch E., Steffen D., Sanders M., Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivy A., Sugano S., White B., Walker D., Woodcock J., Winckler T., Tanaka Y., Shalish G., Schleicher M., Weinscock G., Rosenthal A., Cox E.C., Williams J., Dear P.H., Noegel A.A., Barrett B., Knap A., Nature 0:0-0(2005).

RL "The genome of the social amoeba Dictyostelium discoideum," EMBL; AF46309; AAL/74253.1; -; Genomic_DNA.

DR EMBL; AF46309; AAL/74253.1; -; Genomic_DNA.

DR HSSP; P08716; IMTO.

DR DictyBase; DD80191427; tagd.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0010887; F:ATPase activity; IEA.

DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.

DR GO; GO:0000166; F:nucleotide binding; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004289; F:subtilase activity; IEA.

DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR003593; AAA ATPase.

DR InterPro; IPR011527; ABC membrane 1.

DR InterPro; IPR001140; ABC_TM_transp.

DR InterPro; IPR000209; ABC_transp_like.

DR InterPro; IPR000209; Pept_S8_S53.

DR Pfam; PF00664; ABC_membrane; 1.

DR Pfam; PF00664; ABC_tran; 1.

DR Pfam: PF00082; Peptidase S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR ProDom: PD000006; ABC_transporter; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS50929; ABC_TM1F; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE: PS50893; ABC_TRANSPORTER_2; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 1825 AA; 202642 MW; E28160BC78613A3B CRC64;

Query Match 16.1%; Score 493; DB 2; Length 1825;
Best Local Similarity 24.2%; Pred No. 2,2e-26;
Matches 192; Conservative 105; Mismatches 229; Indels 268; Gaps 31;

```
QY      87 HIXQFNGPILEBT---KQXLEXTG-----AKILDYIPDYAYV---EYEGDVXSX 130
DB      127 YIVQFMDHINDEFREQFQPLINTDIVLDEQPYQSHIVYIPHDSFLVLMNDEQSNLSS 186
QY      131 XXXIPIVSVSVP-----YLPYXIDPOLFTKGASXL-----VK 163
DB      187 KEMVSWIGFEPSPNKLHLYNNEKSIGLPYIKLSDTNSLIQRMENLINSILTSYNSKYK 246
QY      164 AXALPTKQNKKEVQLRGEXIAQXXXS-----NDVYIITAKPEYKMMNDV 208
DB      247 LTLNQKCL-KSIYVQNDSSSSQSCSLVSEKLYQWISEQSESNYIERSEKFOJANRL 305
QY      209 ARGIV-----KADVAQSYGLYGOQIVAVADTGLPTGR---NDS-----SM 247
DB      306 SPKALFGTMDTLVNNDRIDIP-----LRGKGILSLADTGLDGSCHFPDSNNPIPYNSV 360
QY      248 HEAFNGKITLALYALGRTNANDTNHGTIVAGSVLGNKXTN-----KMAPQANLVFQ 300
DB      361 MLNHRKVVYIIGSL--HDNEDYVDGHTVCGSAAAPEDSSLAISFSGLATDAKIAF 418
QY      301 SI-MDSXGGLGSLPBNLQTLFQSAKSAGARITNSMGA---AVNGAYTDSRNVDDYR 355
DB      419 DIASPSNNPEVPPEPDYSQLQPLYNAGRVHGDWSGLSIQCYLGYSYSDAGSIDDFLY 478
QY      356 KN-DWTILFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSF----- 397
DB      479 THPDPIILRAAGNNEYSSLLS-QATAKQVITVGAEGTTHESYTTDALEYSNFETVAKST 537
QY      398 -----GSYAD-----NINHVAQ 409
DB      538 INSLCQSPDKYCTYTTAQCCTEYSTVKGLSGCCTSYKNGSYASIFSSQPELYNENNIC 597
QY      410 PSSRGPFDGRIKPDVMAFGITYILSRSSLA-----PDSFMANHDSKYAYMG 458
DB      598 PSSKGPFDHGRLEPDIVAPGOYITSARNSGANTTDOCGDGLPNTVALISE-----SG 650
QY      459 TSMATPIVAGNVAOLREH-----FVKRGIITPKPSLKALILAGADUXGLGYP-- 506
DB      651 TSMATPLATATTIARQIYLVDSGYTPGSLVSESKLOPTGSLKALMINNAQLINGTFPLS 710
QY      507 -----NGNQGMRVTLDKSLNVAVNESS----- 530
DB      711 STNTNPSNAVFPDTPAGANFVQGMGSLRMEWL---YVESSGYKPKPSRWVGIGELGKDK 767
QY      531 -----XLISTSQKATYXT-----ATAGKP-LKISLVMSDAPASTTASVTLVNDLD 574
DB      768 ASNMKEYSLSLSTQNVSYCFTYPRSSSGSNSGIPRIATLVWTDPPSYSGAKLNLVNNLD 827
QY      575 LVIT-----APNGXYVG-NDPXXKXXXNMDGRNVENVF---INXPOSQTYTIE 620
DB      828 LMTNTTSEFFIYYSNGSGSYNGTKGTTLPLO--DSINNVGIIYTPINTYSEISFPRFI 884
QY      621 VQAYNVVPGPOKFS 634
DB      885 IAGTNIPIGPONFS 898
```

Search completed: April 7, 2006, 17:42:38
Job time : 155.863 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: April 7, 2006, 17:42:57 ; Search time 33.9734 Seconds
(without alignment)
1555.032 Million cell updates/sec

Title: US-10-784-870-1
Perfect score: 3059
Sequence: 1 MEKKKVFSLVSAALISTV.....EVOQNVVPGQXFLATVN 639

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Ispued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/6_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/RB_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3030	99.1	639	2	US-09-509-814A-4
2	3030	99.1	639	2	US-09-920-954-4
3	3032	98.8	640	2	US-09-509-814A-6
4	3022	98.8	640	2	US-09-920-954-6
5	3021	98.8	640	2	US-09-509-814A-8
6	3021	98.8	640	2	US-09-920-954-8
7	3017	98.6	639	2	US-09-509-814A-1
8	3017	98.6	639	2	US-09-920-954-1
9	3007	98.3	640	2	US-09-509-814A-2
10	3007	98.3	640	2	US-09-920-954-2
11	2723	89.0	641	1	US-08-873-479-42
12	2417.5	79.0	635	1	US-08-873-479-43
13	2155	70.4	434	2	US-09-985-689A-1
14	2155	70.4	434	2	US-09-985-689A-2
15	2082	68.1	434	2	US-09-985-689A-6
16	2060.5	67.4	433	2	US-09-985-689A-7
17	1952.5	63.7	433	2	US-09-985-689A-5
18	1948.5	63.7	433	2	US-09-985-689A-3
19	1941.5	63.5	433	2	US-09-985-689A-4
20	1940.5	63.4	433	2	US-09-104-623A-4
21	1940.5	63.4	433	2	US-09-019-532-4
22	1940.5	63.4	433	2	US-09-338-746-4
23	1644.5	53.8	345	2	US-09-512-251A-10
24	1644.5	53.8	345	2	US-09-515-150A-10
25	1644.5	53.8	345	2	US-09-196-281-13
26	1644.5	53.8	345	2	US-10-336-324-10
27	451.5	14.8	659	2	US-08-894-818B-1

28	451.5	14.8	659	2	US-09-445-472-12	Sequence 12, App1
29	451.5	14.8	659	2	US-10-090-624-12	Sequence 12, App1
30	451.5	14.8	659	2	US-09-841-553-1	Sequence 1, App1
31	408	13.3	654	2	US-08-894-818B-35	Sequence 35, App1
32	408	13.3	654	2	US-09-445-472-16	Sequence 16, App1
33	408	13.3	654	2	US-10-090-624-16	Sequence 16, App1
34	408	13.3	654	2	US-08-841-553-35	Sequence 35, App1
35	398	13.0	659	2	US-08-894-818B-5	Sequence 5, App1
36	398	13.0	659	2	US-09-841-553-5	Sequence 5, App1
37	391	12.8	412	2	US-09-445-472-1	Sequence 1, App1
38	391	12.8	412	2	US-10-090-624-1	Sequence 1, App1
39	391	12.8	522	2	US-08-894-818B-3	Sequence 3, App1
40	391	12.8	522	2	US-09-445-472-4	Sequence 4, App1
41	391	12.8	522	2	US-10-090-624-4	Sequence 4, App1
42	391	12.8	522	2	US-09-841-553-3	Sequence 3, App1
43	340	11.1	520	2	US-09-000-016-7	Sequence 7, App1
44	340	11.1	520	2	US-09-514-340-7	Sequence 7, App1
45	340	11.1	734	2	US-09-000-016-4	Sequence 4, App1

ALIGNMENTS

RESULT 1						
US-09-509-814A-4						
Sequence 4, Application US/09509814A						
Patent No. 6376227						
GENERAL INFORMATION:						
APPLICANT: TAKAIWA, MIKIO						
APPLICANT: OKUDA, MITSUYOSHI						
APPLICANT: SAKKI, KATSUHIKA						
APPLICANT: KUBOTA, HIROMI						
APPLICANT: HITOMI, JUN						
APPLICANT: KAGEYAMA, YASUSHI						
APPLICANT: SHIKATA, SHITSUM						
APPLICANT: NOMURA, MASAFUMI						
TITLE OF INVENTION: ALKALINE PROTEASE						
FILE REFERENCE: 0327-0832-0PCT						
CURRENT APPLICATION NUMBER: US/09/509,814A						
CURRENT FILING DATE: 2000-04-06						
PRIOR APPLICATION NUMBER: PCT/JP98/04528						
PRIOR FILING DATE: 1998-10-07						
PRIOR APPLICATION NUMBER: JP 9-274570						
PRIOR FILING DATE: 1997-06-08						
NUMBER OF SEQ ID NOS: 24						
SOFTWARE: Patentin version 3.0						
SEQ ID NO 4						
LENGTH: 639						
TYPE: PRT						
ORGANISM: Bacillus sp.						
US-09-509-814A-4						
Query Match						
Best Local Similarity 93.4%; Score 3030; DB 2; Length 639;						
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;						
QY	1	MEKKKVFSLVSAALISTV	ALXNPSAGARXFDLPFKGIQTITDXXGFSKQXQTGAALF	60		
DB	1	MEKKKVFSLVSAALISTV	ALNPSAGARFDDPFKGIQTITDVSGRSQRTGAALF	60		
QY	61	LVSESVNKLKGLKGLTVPANNCLHIXQFNGPILEBTKQALEXTGAKILDIYIPYAT	120			
DB	61	LVSESVNKLKGLKGLTVPANNCLHIXQFNGPILEBTKQALEXTGAKILDIYIPYAT	120			
QY	121	VEYEGDVAXXXXXIHEVSEVPEYLPYXXIDPOLFTFGASXLVKAXALDPTKQXKEVQLRG	180			
DB	121	VEYEGDVAXXXXXIHEVSEVPEYLPYXXIDPOLFTFGASXLVKAXALDPTKQXKEVQLRG	180			
QY	181	IEIXAQXXXNDVXYITAKPEYKVNDAVGIVKADVAQSSYGLYGQGGQIVAVADTGLDT	240			
DB	181	IEIXAQXXXNDVXYITAKPEYKVNDAVGIVKADVAQSSYGLYGQGGQIVAVADTGLDT	240			
QY	241	GRNDSMEHAFRGKITALLALGRTNNANDTNGHGTIVAGSVLGNKXTKGMAPQANLVQ	300			

DB 241 GRDSSMHEAFRGKITALYALGRNTNANDTNGHGHVAGSVLGNATKNGMAPOANLVQ 300
QY 301 SIMDSXGIGLGLPSNLQTLFSGQXSGAGARIHTNSWGAAVNGAYTTDSRVNDVYRKNDMT 360
DB 301 SIMDSXGIGLGLPSNLQTLFSGQXSGAGARIHTNSWGAAVNGAYTTDSRVNDVYRKNDMT 360
QY 361 ILFAAGNEXPNGGTSAPGTAKNAITVGTATENLRPSFGSYADNINHVAQFSSRGPTKGR 420
DB 361 ILFAAGNEXPNGGTSAPGTAKNAITVGTATENLRPSFGSYADNINHVAQFSSRGPTKGR 420
QY 421 IKPDVMAFGTYILSARSSSLAPDSSFMANHDSKTYAMGTSMAPIVAGNVQOLREHFVN 480
DB 421 IKPDVMAFGTYILSARSSSLAPDSSFMANHDSKTYAMGTSMAPIVAGNVQOLREHFVN 480
QY 481 RGITPKPSLLKALITAGAADXGLGYPNGNGMGVTLTDXSLNVAAYNESSXLSSTOKATY 540
DB 481 RGITPKPSLLKALITAGAADXGLGYPNGNGMGVTLTDXSLNVAAYNESSXLSSTOKATY 540
QY 541 XFTATAGKPKLISLWSDAPASTTASVTLVNDLDTVITAPNGTYVYVNDPFXKPKXNDG 600
DB 541 XFTATAGKPKLISLWSDAPASTTASVTLVNDLDTVITAPNGTYVYVNDPFXKPKXNDG 600
QY 601 RNNVENVFINKPSGTYTIEVOAYNVVPQXFSIATVN 639
DB 601 RNNVENVFINKPSGTYTIEVOAYNVVPQXFSIATVN 639

RESULT 2
US-09-920-954-4
Sequence 4, Application US/0920954
Patent No. 6759228

GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 639
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-920-954-4

Query Match 99.1%; Score 3030; DB 2; Length 639;
Best Local Similarity 93.4%; Pred. No. 1,3e-258;
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 MRKKVFLSVLSAAAILSTVALKNPSAGAXKXFDLDFKGIQTTTDXKGFSGKXQGTGAAP 60
DB 1 MRKKVFLSVLSAAAILSTVALKNPSAGADARTFDLDFKGIQTTTDXKGFSGKXQGTGAAP 60
QY 61 LVSENVTLKXGKLKLETVPANNTLHIXQFNGPILEETKQKLETTGAKIIDIYIPDYAYI 120
DB 61 LVSENVTLKXGKLKLETVPANNTLHIXQFNGPILEETKQKLETTGAKIIDIYIPDYAYI 120
QY 121 VEYEGDVASXXXIETHVESVEPYLPKXIIDPOLFTKGSXLVKAXALDTKQXKEVQARG 180
DB 121 VEYEGDVASXXXIETHVESVEPYLPKXIIDPOLFTKGSXLVKAXALDTKQXKEVQARG 180

DB 121 VEYEGDVASXKRSIETHVESVEPYLPKXIIDPOLFTKGSXLVKAXALDTKQXKEVQARG 180
QY 181 IEXIAQXXXSDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGGCQIYAVADTGLDT 240
DB 181 IEEIQAIVASNDVHYITAKPEYKVMNDVARGIVKADVAQSSYGLYGGCQIYAVADTGLDT 240
QY 241 GRDSSMHEAFRGKITALYALGRNTNANDTNGHGHVAGSVLGNATKNGMAPOANLVQ 300
DB 241 GRDSSMHEAFRGKITALYALGRNTNANDTNGHGHVAGSVLGNATKNGMAPOANLVQ 300
QY 301 SIMDSXGIGLGLPSNLQTLFSGQXSGAGARIHTNSWGAAVNGAYTTDSRVNDVYRKNDMT 360
DB 301 SIMDSXGIGLGLPSNLQTLFSGQXSGAGARIHTNSWGAAVNGAYTTDSRVNDVYRKNDMT 360
QY 361 ILFAAGNEXPNGGTSAPGTAKNAITVGTATENLRPSFGSYADNINHVAQFSSRGPTKGR 420
DB 361 ILFAAGNEXPNGGTSAPGTAKNAITVGTATENLRPSFGSYADNINHVAQFSSRGPTKGR 420
QY 421 IKPDVMAFGTYILSARSSSLAPDSSFMANHDSKTYAMGTSMAPIVAGNVQOLREHFVN 480
DB 421 IKPDVMAFGTYILSARSSSLAPDSSFMANHDSKTYAMGTSMAPIVAGNVQOLREHFVN 480
QY 481 RGITPKPSLLKALITAGAADXGLGYPNGNGMGVTLTDXSLNVAAYNESSXLSSTOKATY 540
DB 481 RGITPKPSLLKALITAGAADXGLGYPNGNGMGVTLTDXSLNVAAYNESSXLSSTOKATY 540
QY 541 XFTATAGKPKLISLWSDAPASTTASVTLVNDLDTVITAPNGTYVYVNDPFXKPKXNDG 600
DB 541 XFTATAGKPKLISLWSDAPASTTASVTLVNDLDTVITAPNGTYVYVNDPFXKPKXNDG 600
QY 601 RNNVENVFINKPSGTYTIEVOAYNVVPQXFSIATVN 639
DB 601 RNNVENVFINKPSGTYTIEVOAYNVVPQXFSIATVN 639

RESULT 3
US-09-509-814A-6
Sequence 6, Application US/09509814A
Patent No. 6376227

GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-509-814A-6

Query Match 98.8%; Score 3022; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 6,5e-258;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 2 RKKVFLSVLSAAAILSTVALKNPSAGAXKXFDLDFKGIQTTTDXKGFSGKXQGTGAAP 61
DB 2 RKKVFLSVLSAAAILSTVALKNPSAGAXKXFDLDFKGIQTTTDXKGFSGKXQGTGAAP 61
QY 62 VESNVTLKXGKLKLETVPANNTLHIXQFNGPILEETKQKLETTGAKIIDIYIPDYAYI 121
DB 62 VESNVTLKXGKLKLETVPANNTLHIXQFNGPILEETKQKLETTGAKIIDIYIPDYAYI 121

```
Db 63 VESENVKLPKGLQKLETVPANCKLHIIQFNGPIIETQOLEKTAKLIDYIPDYAYIV 122
Qy 122 EYEGDVXSXXXXIEHVESVEPYLPYXXIDPOLFTKGASLYKAXALDTQXNKEVOLRGI 181
Db 123 EYEGDVKSATSTIEHVESVEPYLPYIRIDPOLFTKGASLYKAXALDTQXNKEVOLRGI 182
Qy 182 EXIAOXXXSNDVXYITTAKEPYKMDVARGIKADVAQSSYGLYGGQIIVAADTGLDYG 241
Db 183 EOIAPALSNDVLYTTAKPEYKMDVARGIKADVAQSSYGLYGGQIIVAADTGLDYG 242
Qy 242 RUDSSMHEAFRGKITLVALGRTNNANDNGHGHVAGSVLGNSTNKGMAPQANLVPOS 301
Db 243 RUDSSMHEAFRGKITLVALGRTNNANDNGHGHVAGSVLGNSTNKGMAPQANLVPOS 302
Qy 302 IMDSXGGLGGLPSNLTFLSQXVSAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDMTI 361
Db 303 IMDSXGGLGGLPSNLTFLSQXVSAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDMTI 362
Qy 362 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDGR 421
Db 363 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDGR 422
Qy 422 KPDVAPGTXYILSARSLAPDSSFMANHDSKTYAMGTSMAPIYAGNVAOLREHFVKR 481
Db 423 KPDVAPGTXYILSARSLAPDSSFMANHDSKTYAMGTSMAPIYAGNVAOLREHFVKR 482
Qy 482 GITPRSLKALILGADADIGLGYPNGNGMGWRTLDKSLNVAAYNESSSLSTSQATYS 541
Db 483 GITPRSLKALILGADADIGLGYPNGNGMGWRTLDKSLNVAAYNESSSLSTSQATYS 542
Qy 542 PTATAGKPLKISLVMSDAPASTTASVTLVNDLVTITAPNGTYVYVNDYXXPKXNDGR 601
Db 543 PTATAGKPLKISLVMSDAPASTTASVTLVNDLVTITAPNGTYVYVNDYXXPKXNDGR 602
Qy 602 NNVENVFINKXPOSQTYTIEVQAYNVPGPQXSLAIYN 639
Db 603 NNVENVFINKXPOSQTYTIEVQAYNVPGPQXSLAIYN 640
```

RESULT 4

```
US-09-920-954-6
; Sequence 6, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 6
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-920-954-6
```

```
Query Match 98.8%; Score 3022; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 6.5e-258;
```

```
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
Qy 2 RKKKVFSLSTLSAAALISTALKNPSAGAKAXDDLPKGIQTTTDXKGSKQOTGAATL 61
Db 3 KKKKVFSLSTLSAAALISTALSNPSAGAKARNDDLPKGIQTTTDXKGSKQOTGAATL 62
Qy 62 VESENVKLPKGLQKLETVPANCKLHIIQFNGPIIETQOLEKTAKLIDYIPDYAYIV 121
Db 63 VESENVKLPKGLQKLETVPANCKLHIIQFNGPIIETQOLEKTAKLIDYIPDYAYIV 122
Qy 122 EYEGDVXSXXXXIEHVESVEPYLPYXXIDPOLFTKGASLYKAXALDTQXNKEVOLRGI 181
Db 123 EYEGDVKSATSTIEHVESVEPYLPYIRIDPOLFTKGASLYKAXALDTQXNKEVOLRGI 182
Qy 182 EXIAOXXXSNDVXYITTAKEPYKMDVARGIKADVAQSSYGLYGGQIIVAADTGLDYG 241
Db 183 EOIAPALSNDVLYTTAKPEYKMDVARGIKADVAQSSYGLYGGQIIVAADTGLDYG 242
Qy 242 RUDSSMHEAFRGKITLVALGRTNNANDNGHGHVAGSVLGNSTNKGMAPQANLVPOS 301
Db 243 RUDSSMHEAFRGKITLVALGRTNNANDNGHGHVAGSVLGNSTNKGMAPQANLVPOS 302
Qy 302 IMDSXGGLGGLPSNLTFLSQXVSAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDMTI 361
Db 303 IMDSXGGLGGLPSNLTFLSQXVSAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDMTI 362
Qy 362 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDGR 421
Db 363 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDGR 422
Qy 422 KPDVAPGTXYILSARSLAPDSSFMANHDSKTYAMGTSMAPIYAGNVAOLREHFVKR 481
Db 423 KPDVAPGTXYILSARSLAPDSSFMANHDSKTYAMGTSMAPIYAGNVAOLREHFVKR 482
Qy 482 GITPRSLKALILGADADIGLGYPNGNGMGWRTLDKSLNVAAYNESSSLSTSQATYS 541
Db 483 GITPRSLKALILGADADIGLGYPNGNGMGWRTLDKSLNVAAYNESSSLSTSQATYS 542
Qy 542 PTATAGKPLKISLVMSDAPASTTASVTLVNDLVTITAPNGTYVYVNDYXXPKXNDGR 601
Db 543 PTATAGKPLKISLVMSDAPASTTASVTLVNDLVTITAPNGTYVYVNDYXXPKXNDGR 602
Qy 602 NNVENVFINKXPOSQTYTIEVQAYNVPGPQXSLAIYN 639
Db 603 NNVENVFINKXPOSQTYTIEVQAYNVPGPQXSLAIYN 640
```

RESULT 5

```
US-09-509-814A-8
; Sequence 8, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
```

ORGANISM: Bacillus sp.
US-09-509-814A-8

Query Match 98.8%; Score 3021; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 8e-258;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```
OY 2 RKKKVFLSVLSAAAILSTVALXNPSAGAKRFPDLDFKGIQTTTDXGSKQKQGTGAFL 61
DB 3 KKKKVFSLVLSAAAILSTVALSNPSAGARNFDLPFKGIQTTTDXGSKQKQGTGAFL 62
OY 62 VESENVKLXKGLXKKLETPANNKHLHXQFNGLIETKQLEXTGAKILDIYIPDAYIV 121
DB 63 VESENVKLXKGLXKKLETPANNKHLHXQFNGLIETKQLEXTGAKILDIYIPDAYIV 122
OY 122 EYEGDVASXXXXIHHVESVEPYLPYXIDPOLFTKGASXLVKAXALDTKQNKKEVOLRGI 181
DB 123 EYEGDVASATSTIEHVESVEPYLPYRIDPOLFTKGASELVKAXALDTKQNKKEVOLRGI 182
OY 182 EXIAQXXXNDVXYITAKPEYKMNDDVARGIVKADVAOSSYGLYGQGIIVAVADTGLDTG 241
DB 183 EQLAQFALSNDVLYITAKPEYKMNDDVARGIVKADVAOSSYGLYGQGIIVAVADTGLDTG 242
OY 242 RNDSSMHEAFRGKITLALYALGRTNNANDTNGHGVASVLGNGXTKGMAPQANLVFQS 301
DB 243 RNDSSMHEAFRGKITLALYALGRTNNANDTNGHGVASVLGNGXTKGMAPQANLVFQS 302
OY 302 IMDSXGGLGGLPSNLQTLFSGQXSGAGARIHTNSWGAANGAYTTDSRVVDYVRKNDMTI 361
DB 303 IMDSGGLGGLPSNLQTLFSGQYSGAGARIHTNSWGAANGAYTTDSRVVDYVRKNDMTI 362
OY 362 LPAAGNEKPNCGTTSAPRTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKGR 421
DB 363 LPAAGNEKPNCGTTSAPRTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKGR 422
OY 422 KPDVMAPEGTXILSARSSLAPDSSFMANHDSKYAYNGTSMATPIYAGNVAOLRHFVNKR 481
DB 423 KPDVMAPEGTXILSARSSLAPDSSFMANHDSKYAYNGTSMATPIYAGNVAOLRHFVNKR 482
OY 482 GITPKPSLLKALJAGAADXGLGYPNGQGWGRVTLDSLNVAAYNESSSLSTSQKATYS 541
DB 483 GITPKPSLLKALJAGAADXGLGYPNGQGWGRVTLDSLNVAAYNESSSLSTSQKATYS 542
OY 542 FTATAGKPLKLSLVWSDAPASTTASVTLVNDLDVITAPNGTXVYGNDPXXPKXNMGR 601
DB 543 FTATAGKPLKLSLVWSDAPASTTASVTLVNDLDVITAPNGTXVYGNDPXXPKXNMGR 602
OY 602 NNVENVFINKPQSGTYTIEVQAYNVPGPQKFSLAIVN 639
DB 603 NNVENVFINKPQSGTYTIEVQAYNVPGPQKFSLAIVN 640
```

RESULT 6
US-09-920-954-8
Sequence 8, Application US/09920954
Patent No. 6759228
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent In version 3.0
SEQ ID NO 8
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-920-954-8

Query Match 98.8%; Score 3021; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 8e-258;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```
OY 2 RKKKVFLSVLSAAAILSTVALXNPSAGAKRFPDLDFKGIQTTTDXGSKQKQGTGAFL 61
DB 3 KKKKVFSLVLSAAAILSTVALSNPSAGARNFDLPFKGIQTTTDXGSKQKQGTGAFL 62
OY 62 VESENVKLXKGLXKKLETPANNKHLHXQFNGLIETKQLEXTGAKILDIYIPDAYIV 121
DB 63 VESENVKLXKGLXKKLETPANNKHLHXQFNGLIETKQLEXTGAKILDIYIPDAYIV 122
OY 122 EYEGDVASXXXXIHHVESVEPYLPYXIDPOLFTKGASXLVKAXALDTKQNKKEVOLRGI 181
DB 123 EYEGDVASATSTIEHVESVEPYLPYRIDPOLFTKGASELVKAXALDTKQNKKEVOLRGI 182
OY 182 EXIAQXXXNDVXYITAKPEYKMNDDVARGIVKADVAOSSYGLYGQGIIVAVADTGLDTG 241
DB 183 EQLAQFALSNDVLYITAKPEYKMNDDVARGIVKADVAOSSYGLYGQGIIVAVADTGLDTG 242
OY 242 RNDSSMHEAFRGKITLALYALGRTNNANDTNGHGVASVLGNGXTKGMAPQANLVFQS 301
DB 243 RNDSSMHEAFRGKITLALYALGRTNNANDTNGHGVASVLGNGXTKGMAPQANLVFQS 302
OY 302 IMDSXGGLGGLPSNLQTLFSGQXSGAGARIHTNSWGAANGAYTTDSRVVDYVRKNDMTI 361
DB 303 IMDSGGLGGLPSNLQTLFSGQYSGAGARIHTNSWGAANGAYTTDSRVVDYVRKNDMTI 362
OY 362 LPAAGNEKPNCGTTSAPRTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKGR 421
DB 363 LPAAGNEKPNCGTTSAPRTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKGR 422
OY 422 KPDVMAPEGTXILSARSSLAPDSSFMANHDSKYAYNGTSMATPIYAGNVAOLRHFVNKR 481
DB 423 KPDVMAPEGTXILSARSSLAPDSSFMANHDSKYAYNGTSMATPIYAGNVAOLRHFVNKR 482
OY 482 GITPKPSLLKALJAGAADXGLGYPNGQGWGRVTLDSLNVAAYNESSSLSTSQKATYS 541
DB 483 GITPKPSLLKALJAGAADXGLGYPNGQGWGRVTLDSLNVAAYNESSSLSTSQKATYS 542
OY 542 FTATAGKPLKLSLVWSDAPASTTASVTLVNDLDVITAPNGTXVYGNDPXXPKXNMGR 601
DB 543 FTATAGKPLKLSLVWSDAPASTTASVTLVNDLDVITAPNGTXVYGNDPXXPKXNMGR 602
OY 602 NNVENVFINKPQSGTYTIEVQAYNVPGPQKFSLAIVN 639
DB 603 NNVENVFINKPQSGTYTIEVQAYNVPGPQKFSLAIVN 640
```

RESULT 7
US-09-509-814A-1
Sequence 1, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE

```
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPE: PRF
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match      98.6%; Score 3017; DB 2; Length 639;
Best Local Similarity 100.0%; Pred. No. 1,88-257;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
QY 1 MRKKVPLSVLSAAAILSTVALXNPSAGARXFDLDFKGIOTTTDXKGSFKOXOTGAAP 60
DB 1 MRKKVPLSVLSAAAILSTVALXNPSAGARXFDLDFKGIOTTTDXKGSFKOXOTGAAP 60
QY 61 LVESENVLKLKXGLKXKLETPVANNKLHIKQFNGPILLETQKXILEXTGAKILDIYIPDVAI 120
DB 61 LVESENVLKLKXGLKXKLETPVANNKLHIKQFNGPILLETQKXILEXTGAKILDIYIPDVAI 120
QY 121 VEYEDDVXSKXXXIHEVESVEPYLDYXXIDPOLFTKGASXLVKAYALDTKXNKEVOLRG 180
DB 121 VEYEDDVXSKXXXIHEVESVEPYLDYXXIDPOLFTKGASXLVKAYALDTKXNKEVOLRG 180
QY 181 IEXIQXXXNDVXXITAKPEYKWNNDVARGIVKADVQSSYGLVGQGIIVAVADTGLDT 240
DB 181 IEXIQXXXNDVXXITAKPEYKWNNDVARGIVKADVQSSYGLVGQGIIVAVADTGLDT 240
QY 241 GRNDSMHEAFRGKITLALYALGRTNNANDTNGHGTIVAGSVLGNKXTKGMAPQANLVFQ 300
DB 241 GRNDSMHEAFRGKITLALYALGRTNNANDTNGHGTIVAGSVLGNKXTKGMAPQANLVFQ 300
QY 301 SIMDSXGGLGGLPSNLQTLFSQAXSAGARIHTNSWGAAVNGAYTTDSRNVDDYERKNDMT 360
DB 301 SIMDSXGGLGGLPSNLQTLFSQAXSAGARIHTNSWGAAVNGAYTTDSRNVDDYERKNDMT 360
QY 361 ILFAGNEXPNNGGTISAPGTAKNALTGVATENLRPSFGSYADNINHVAOFSSRGPTKGR 420
DB 361 ILFAGNEXPNNGGTISAPGTAKNALTGVATENLRPSFGSYADNINHVAOFSSRGPTKGR 420
QY 421 IKPDVMAFGTXILSARSSILAPDSSFWMANHDSKAYMGGTSMATPIVAGNVALRSHFVYN 480
DB 421 IKPDVMAFGTXILSARSSILAPDSSFWMANHDSKAYMGGTSMATPIVAGNVALRSHFVYN 480
QY 481 RGITPKPSLLKALITAGADKGLGYPNGQGWGRVTLDKSLNAVAVYNESSXISTQKATY 540
DB 481 RGITPKPSLLKALITAGADKGLGYPNGQGWGRVTLDKSLNAVAVYNESSXISTQKATY 540
QY 541 XFTATAGPKIKTSLWSDAPASTASVTLVNDLIVTPNCTXYVGVNDPFXKXXXNDMG 600
DB 541 XFTATAGPKIKTSLWSDAPASTASVTLVNDLIVTPNCTXYVGVNDPFXKXXXNDMG 600
QY 601 RNNVENVPINXPOSGTYTIEVQAVNVPVPOXFSLAIYN 639
DB 601 RNNVENVPINXPOSGTYTIEVQAVNVPVPOXFSLAIYN 639

RESULT 8
US-09-920-954-1
; Sequence 1, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEBT, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 639
; TYPE: PRT
```

```
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
```

```
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-920-954-1
```

Query Match 98.6%; Score 3017; DB 2; Length 639;
Best Local Similarity 100.0%; Pred. No. 1.8e-257;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRKKVPLSVLSAAAILSTVALXNPSAGAKRPDIDFKGIQTTTDXKXGSKXQTGAANF 60
DB 1 MRKKVPLSVLSAAAILSTVALXNPSAGAKRPDIDFKGIQTTTDXKXGSKXQTGAANF 60
QY 61 LVSESNVQLXKXGKGLSTVPANNKLHIQFNGPILEBTKOXLKXTGAAILDIPIDYAYI 120
DB 61 LVSESNVQLXKXGKGLSTVPANNKLHIQFNGPILEBTKOXLKXTGAAILDIPIDYAYI 120
QY 121 VEYBGDVXSKXXKXIEHVSVEBPYLPXYXIDPOLFTKGSXLVKAXALDTKQXNKEVOLRG 180
DB 121 VEYBGDVXSKXXKXIEHVSVEBPYLPXYXIDPOLFTKGSXLVKAXALDTKQXNKEVOLRG 180
```

```
QY 181 IEXIAOXXXXNDVXYITLAKBEYKVMNDVARGIVKADVAGSSYGLVGQGOI VAVADTGLPT 240
DB 181 IEXIAOXXXXNDVXYITLAKBEYKVMNDVARGIVKADVAGSSYGLVGQGOI VAVADTGLPT 240
QY 241 GRNDSMHEAFRGKITLALYALGRTNANDJNGHGTIVAGSVLGNXTNNGMAPOANLVFO 300
DB 241 GRNDSMHEAFRGKITLALYALGRTNANDJNGHGTIVAGSVLGNXTNNGMAPOANLVFO 300
QY 301 SIMDSXGIGLGLPSULQTLPSQAKSAGARIHTNSGAAVNGAYTTDSRRVDYVRKNDMT 360
DB 301 SIMDSXGIGLGLPSULQTLPSQAKSAGARIHTNSGAAVNGAYTTDSRRVDYVRKNDMT 360
QY 361 ILFAAGNEXPNGGTSAPGTAKNALTVAATEMLRPSFGSYADINHHVAPFSRGPTKDR 420
DB 361 ILFAAGNEXPNGGTSAPGTAKNALTVAATEMLRPSFGSYADINHHVAPFSRGPTKDR 420
QY 421 IKPDVMAFGTXILSARSSILAPDSFWANHDSKYAYMGTSMTPTVAGVNAOLREHFVN 480
DB 421 IKPDVMAFGTXILSARSSILAPDSFWANHDSKYAYMGTSMTPTVAGVNAOLREHFVN 480
QY 481 RGITRPSILKALITAGAADKGLYPNGNGKGRVTLDKSLNAVYNESSXLSQKATY 540
DB 481 RGITRPSILKALITAGAADKGLYPNGNGKGRVTLDKSLNAVYNESSXLSQKATY 540
QY 541 XFTATAGRPKISLWSDAPASTTASVTLVNDLDTVITAPNCTXYVGNDFXXPKXNMWG 600
DB 541 XFTATAGRPKISLWSDAPASTTASVTLVNDLDTVITAPNCTXYVGNDFXXPKXNMWG 600
QY 601 RNNVENFPLNKPSGTYTIEVQAVNVPVGPQXSLAIYN 639
DB 601 RNNVENFPLNKPSGTYTIEVQAVNVPVGPQXSLAIYN 639
```

```
RESULT 9
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASR
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
```

NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature

LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-2

Query Match 98.3%; Score 3007; DB 2; Length 640;
Best Local Similarity 100.0%; Pred. No. 1,4e-256;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KKKVFLSVLSAAALISTVALKNPSAGXARXFPDDLFRKGIQTITDXXGSKXQGTAAAFV 62
DB 4 KKKVFLSVLSAAALISTVALKNPSAGXARXFPDDLFRKGIQTITDXXGSKXQGTAAAFV 63
QY ESENVKXKXGKXKGLKLETPANNKXKHIKQFNGLIETKXKLEKXGKILDIYIPYAYIV 122
DB 64 ESENVKXKXGKXKGLKLETPANNKXKHIKQFNGLIETKXKLEKXGKILDIYIPYAYIV 123
QY 123 YEGDVXSKXXXIEHVSVEPYLPYXIDPOLFTKGASXLVKAALDTKXNKEVQLGIE 182
DB 124 YEGDVXSKXXXIEHVSVEPYLPYXIDPOLFTKGASXLVKAALDTKXNKEVQLGIE 183
QY 183 XIAXXXXSNDVXYITTAPEYKVMNDVARGIKADVAOSSYGLYQGGQIVAVADTGLTGR 242
DB 184 XIAXXXXSNDVXYITTAPEYKVMNDVARGIKADVAOSSYGLYQGGQIVAVADTGLTGR 243
QY 243 NDSMEHAFRGKITLALAGRTNANDNGHGVAGSYGNGKTNKGMAPOANLVPOSI 302
DB 244 NDSMEHAFRGKITLALAGRTNANDNGHGVAGSYGNGKTNKGMAPOANLVPOSI 303
QY 303 MDSXGGLGGLPSNLQTLFSQXKSGARIHNTSWGAAVNGAYTTDSRVVDVYVRKNDWTIL 362
DB 304 MDSXGGLGGLPSNLQTLFSQXKSGARIHNTSWGAAVNGAYTTDSRVVDVYVRKNDWTIL 363

NAME/KEY: misc feature	LOCATION: (71)-(71)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (75)-(75)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (90)-(90)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (103)-(103)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (106)-(106)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (129)-(129)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (131)-(131)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (132)-(132)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (133)-(133)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (134)-(134)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (147)-(147)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (149)-(149)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (161)-(161)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (166)-(166)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (173)-(173)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (184)-(184)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (188)-(188)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (189)-(189)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (190)-(190)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (195)-(195)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (287)-(287)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (307)-(307)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (325)-(325)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature	LOCATION: (370)-(370)	
OTHER INFORMATION: Xaa	is any amino acid	
NAME/KEY: misc feature		

```
LOCATION: (432) .. (432)
NAME/KEY: misc feature
LOCATION: (502) .. (502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (532) .. (532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (542) .. (542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (585) .. (585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592) .. (592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (593) .. (593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595) .. (595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596) .. (596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (597) .. (597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (612) .. (612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (633) .. (633)
OTHER INFORMATION: Xaa is any amino acid
US-09-920-954-2
```

Query Match 98.3%; Score 3007; DB 2; Length 640;

Best Local Similarity 100.0%; Pred. No. 1.4e-256; Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
3 KKKVFLSVLSAAALISTVALXNPSAGXARXFDLDFKGIOTTTDXGFSKQXGTGAAPLV 62
4 KKKVFLSVLSAAALISTVALXNPSAGXARXFDLDFKGIOTTTDXGFSKQXGTGAAPLV 63
63 ESENVKLLKGLKLETPVANNKLIHQNGPILBERTKQXLEXTAKILDIYIPDYAIVE 122
64 ESENVKLLKGLKLETPVANNKLIHQNGPILBERTKQXLEXTAKILDIYIPDYAIVE 123
123 YEGDVXSKXXKIEHVESVEPYLPXYXIDPQLFTKGASXLVKAXALDTKQXNKEVOLRGIE 182
124 YEGDVXSKXXKIEHVESVEPYLPXYXIDPQLFTKGASXLVKAXALDTKQXNKEVOLRGIE 183
183 XIAQXXKXNDVYITAKPEYKVMNDVARGIVADVAQSSYGLYGQGIIVAVADTGLDTR 242
184 XIAQXXKXNDVYITAKPEYKVMNDVARGIVADVAQSSYGLYGQGIIVAVADTGLDTR 243
243 NDSNHEAFRGKITLTLVLTGRTNANDPMTGHTHVASLTLGNGKTNKGMAPQANLVPSI 302
244 NDSNHEAFRGKITLTLVLTGRTNANDPMTGHTHVASLTLGNGKTNKGMAPQANLVPSI 303
303 MDSXGGLGGLPSNLQTLFSAQXASAGARLITNSMGAAVNGAYTTDSRNVDDVYRKNDMTIL 362
304 MDSXGGLGGLPSNLQTLFSAQXASAGARLITNSMGAAVNGAYTTDSRNVDDVYRKNDMTIL 363
363 FAAGNEXNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHAQSSRSRPTDGRIG 422
364 FAAGNEXNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHAQSSRSRPTDGRIG 423
423 PDVMAFGTXILISARSLAPDSSFMNNDHDKVAYMGTSMAATPIVAGNVAAQLREHVVXRG 482
424 PDVMAFGTXILISARSLAPDSSFMNNDHDKVAYMGTSMAATPIVAGNVAAQLREHVVXRG 483
```

```
483 ITPKSLKALINAGAADXGYPNGQMGRTLDKSLNAYVNSSXLSTSQATYXF 542
484 ITPKSLKALINAGAADXGYPNGQMGRTLDKSLNAYVNSSXLSTSQATYXF 543
543 TATAGKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTYVGNDFXKXXNMDGRN 602
544 TATAGKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTYVGNDFXKXXNMDGRN 603
603 NVENVFIXXPOSGTYTIEVOQAYNVVGPQXPSLAIYN 639
604 NVENVFIXXPOSGTYTIEVOQAYNVVGPQXPSLAIYN 640
```

RESULT 11

US-08-873-479-42

```
/ Sequence 42, Application US/08873479
/ Patent No. 5891701
/ GENERAL INFORMATION:
/ APPLICANT: Symyx, Alan
/ TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
/ TITLE OF INVENTION: Having Protease Activity
/ NUMBER OF SEQUENCES: 57
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: No. 5891701 of No. 5891701 of No. 5891701 of America
/ STREET: 405 Lexington Avenue
/ CITY: New York
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10174
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTA for windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/873, 479
/ FILING DATE: 12-JUN-1997
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Agis, Cheryl H
/ REGISTRATION NUMBER: 34,086
/ REFERENCE/DOCKET NUMBER: 5251,000-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-867-0123
/ TELEFAX: 212-878-9655
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 42:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 641 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FRAGMENT TYPE: internal
/ US-08-873-479-42
```

Query Match 89.0%; Score 2723; DB 1; Length 641;

Best Local Similarity 82.9%; Pred. No. 1.5e-231; Matches 532; Conservative 40; Mismatches 66; Indels 4; Gaps 2;

```
1 MRK---KKVFLSVLSAAALISTVALXNPSAGXARXFDLDFKGIOTTTDXGFSKQXGTGA 57
2 MRK---KKVFLSVLSAAALISTVALXNPSAGXARXFDLDFKGIOTTTDXGFSKQXGTGA 58
3 MRK---KKVFLSVLSAAALISTVALXNPSAGXARXFDLDFKGIOTTTDXGFSKQXGTGA 59
4 MRK---KKVFLSVLSAAALISTVALXNPSAGXARXFDLDFKGIOTTTDXGFSKQXGTGA 60
58 AAPVSENVKLLKGLKLETPVANNKLIHQNGPILBERTKQXLEXTAKILDIYIPDY 117
61 ASPVSENVKLLKGLKLETPVANNKLIHQNGPILBERTKQXLEXTAKILDIYIPDY 120
118 AYIYVEGDVXSKXXKIEHVESVEPYLPXYXIDPQLFTKGASXLVKAXALDTKQXNKEVO 177
121 AYIYVEGDVXSKXXKIEHVESVEPYLPXYXIDPQLFTKGASXLVKAXALDTKQXNKEVO 180
```


SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-1

Query Match 70.4%; Score 2155; DB 2; Length 434;
Best Local Similarity 96.3%; Pred. No. 1.1e-181;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 206 NDVARGIVKADVAQSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALVAGRTN 265
DB 1 NDVARGIVKADVAQSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALVAGRTN 60
QY 266 NANTNGHGTIVAGSVLGNKTKNGMAPQANLVFQSIMDSXGGLGSLPSNIQTLFSGQXS 325
DB 61 NANTNGHGTIVAGSVLGNKTKNGMAPQANLVFQSIMDSXGGLGSLPSNIQTLFSGQXS 120
QY 326 AGARHTNSMGAANVGAATTTDSRNVDPVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
DB 121 AGARHTNSMGAANVGAATTTDSRNVDPVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
QY 386 TVGATENLRPSFGSYADININVAQFSSRGPTKGRIKPDVMAPTXILSARSILAPDSSE 445
DB 181 TVGATENLRPSFGSYADININVAQFSSRGPTKGRIKPDVMAPTXILSARSILAPDSSE 240
QY 446 WANHDSKAYVGGTSMATPIVAGNVAQLREHFVNKRGITTPKPSILKALIGAADVIGY 505
DB 241 WANHDSKAYVGGTSMATPIVAGNVAQLREHFVNKRGITTPKPSILKALIGAADVIGY 300
QY 506 PNGQGMGRVTLDSKLVAAVYNESSXLSSTOKATYXFTATGKPLKISLVMSDAPASTTA 565
DB 301 PNGQGMGRVTLDSKLVAAVYNESSXLSSTOKATYXFTATGKPLKISLVMSDAPASTTA 360
QY 566 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXKXNDGRNVENVFINXPOSQTYTIEVQAYN 625
DB 361 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXKXNDGRNVENVFINXPOSQTYTIEVQAYN 420
QY 626 VPVGPOKFSIAIYN 639
DB 421 VPVGPOKFSIAIYN 434

RESULT 14

US-09-985-689A-2
Sequence 2, Application US/09985689A
Patent No. 6803222
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAMA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483USO
CURRENT APPLICATION NUMBER: US/09/985, 689A
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin version 3.1
SEQ ID NO 2
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-2

Query Match 70.4%; Score 2155; DB 2; Length 434;
Best Local Similarity 96.3%; Pred. No. 1.1e-181;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 206 NDVARGIVKADVAQSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALVAGRTN 265
DB 1 NDVARGIVKADVAQSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALVAGRTN 60
QY 266 NANTNGHGTIVAGSVLGNKTKNGMAPQANLVFQSIMDSXGGLGSLPSNIQTLFSGQXS 325
DB 61 NANTNGHGTIVAGSVLGNKTKNGMAPQANLVFQSIMDSXGGLGSLPSNIQTLFSGQXS 120
QY 326 AGARHTNSMGAANVGAATTTDSRNVDPVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
DB 121 AGARHTNSMGAANVGAATTTDSRNVDPVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
QY 386 TVGATENLRPSFGSYADININVAQFSSRGPTKGRIKPDVMAPTXILSARSILAPDSSE 445
DB 181 TVGATENLRPSFGSYADININVAQFSSRGPTKGRIKPDVMAPTXILSARSILAPDSSE 240
QY 446 WANHDSKAYVGGTSMATPIVAGNVAQLREHFVNKRGITTPKPSILKALIGAADVIGY 505
DB 241 WANHDSKAYVGGTSMATPIVAGNVAQLREHFVNKRGITTPKPSILKALIGAADVIGY 300
QY 506 PNGQGMGRVTLDSKLVAAVYNESSXLSSTOKATYXFTATGKPLKISLVMSDAPASTTA 565
DB 301 PNGQGMGRVTLDSKLVAAVYNESSXLSSTOKATYXFTATGKPLKISLVMSDAPASTTA 360
QY 566 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXKXNDGRNVENVFINXPOSQTYTIEVQAYN 625
DB 361 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXKXNDGRNVENVFINXPOSQTYTIEVQAYN 420
QY 626 VPVGPOKFSIAIYN 639
DB 421 VPVGPOKFSIAIYN 434

RESULT 15

US-09-985-689A-6
Sequence 6, Application US/09985689A
Patent No. 6803222
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAMA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483USO
CURRENT APPLICATION NUMBER: US/09/985, 689A
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin version 3.1
SEQ ID NO 6
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-6

Query Match 68.1%; Score 2082; DB 2; Length 434;
Best Local Similarity 91.5%; Pred. No. 3e-175;
Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 206 NDVARGIVKADVAQSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALVAGRTN 265
DB 1 NDVARGIVKADVAQSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALVAGRTN 60


```

Qy 266 NANDTRGHGTHVAGSVLGNXITKGMAPQANLVFQSIINDSXGGLGGLPSNLQTLFQOAXS 325
Db 61 NANDPRGHGTHVAGSVLGNXITKGMAPQANLVFQSVMDXNGGLGGLPSNVSTLFSQAYS 120
Qy 326 AGARITHNSMGAUVNCAVTTDSRNVDDYRKIDMTILPAAGNEXPNGGTISAPGTAKNAI 385
Db 121 AGARITHNSMGAUVNCAVTTDSRNVDDYRKIDMAVLPAAGNEGPNGGTISAPGTAKNAI 180
Qy 386 TVGATENTLRPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAFGTYILSARSSLAEDSSF 445
Db 181 TVGATENTLRPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAFGTYILSARSSLAEDSSF 240
Qy 446 WANNHDSKYAYMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKALIAQAADXGLGY 505
Db 241 WANNHDSKYAYMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKALIAQAADXGLGY 300
Qy 506 PNGNOQMGHVTLDKSLNVAVYNESSXLSSTOKATYXFTATAGKPLKISLVMSDAPASTTA 565
Db 301 PNGNOQMGHVTLDKSLNVAVYNESSXLSSTOKATYXFTATAGKPLKISLVMSDAPASTTA 360
Qy 566 SVTLVNDLDELVTAPNGTYVGNDFPKXPXXNMWGRNVEENVFINXPOSQTYTIEVOAYN 625
Db 361 SVTLVNDLDELVTAPNGTYVGNDFPKXPXXNMWGRNVEENVFINXPOSQTYTIEVOAYN 420
Qy 626 VPVGFQXPSLAIVN 639
Db 421 VPVGFQXPSLAIVN 434

```

Search completed: April 7, 2006, 17:44:48
 Job time : 35.9734 secs

;

;

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using BW model

Run on: April 7, 2006, 17:58:27 ; Search time 118.407 Seconds

(without alignments)
2254.868 Million cell updates/sec

Title: US-10-784-870-1

Perfect score: 3059
Sequence: 1 MRKKVPLSVLSAAAILSTV.....EVOAYNVPGVQXPSLAIYN 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3030	99.1	639	3	US-09-920-954-4
2	3030	99.1	639	4	US-10-784-870-4
3	3022	98.8	640	3	US-09-920-954-6
4	3022	98.8	640	4	US-10-456-479-4*
5	3022	98.8	640	4	US-10-784-870-6
6	3022	98.8	640	5	US-10-820-712A-3*
7	3022	98.8	640	5	US-10-820-712A-3*
8	3021	98.8	640	3	US-09-920-954-8
9	3021	98.8	640	4	US-10-784-870-8
10	3017	98.6	639	3	US-09-920-954-1
11	3017	98.6	639	4	US-10-784-870-1
12	3007	98.3	640	4	US-09-920-954-2
13	3007	98.3	640	4	US-10-784-870-2
14	2155	70.4	434	3	US-09-985-689A-1
15	2155	70.4	434	3	US-09-985-689A-2
16	2155	70.4	434	4	US-10-385-662-2
17	2155	70.4	434	4	US-10-456-479-2
18	2155	70.4	434	4	US-10-456-479-10
19	2155	70.4	434	4	US-10-456-479-11
20	2155	70.4	434	4	US-10-837-566-1
21	2155	70.4	434	4	US-10-837-566-2
22	2155	70.4	434	5	US-10-820-712A-1
23	2155	70.4	434	5	US-10-820-712A-12
24	2155	70.4	434	5	US-10-820-712A-14
25	2155	70.4	434	5	US-10-820-714A-1
26	2155	70.4	434	5	US-10-820-714A-13
27	2155	70.4	434	5	US-10-820-714A-15

28	2082	68.1	434	3	US-09-985-689A-6	Sequence 6, App1
29	2082	68.1	434	4	US-10-456-479-15	Sequence 15, App1
30	2082	68.1	434	4	US-10-837-566-6	Sequence 6, App1
31	2082	68.1	434	5	US-10-820-712A-22	Sequence 22, App1
32	2082	68.1	434	5	US-10-820-714A-23	Sequence 23, App1
33	2060.5	67.4	433	3	US-09-985-689A-7	Sequence 7, App1
34	2060.5	67.4	433	4	US-10-456-479-16	Sequence 16, App1
35	2060.5	67.4	433	4	US-10-837-566-7	Sequence 7, App1
36	2060.5	67.4	433	5	US-10-820-712A-23	Sequence 23, App1
37	2060.5	67.4	433	5	US-10-820-714A-24	Sequence 24, App1
38	1952.5	63.8	433	3	US-09-985-689A-5	Sequence 5, App1
39	1952.5	63.8	433	4	US-10-456-479-14	Sequence 14, App1
40	1952.5	63.8	433	4	US-10-837-566-5	Sequence 5, App1
41	1952.5	63.8	433	5	US-10-820-712A-20	Sequence 20, App1
42	1952.5	63.8	433	5	US-10-820-714A-21	Sequence 21, App1
43	1948.5	63.7	433	3	US-09-985-689A-3	Sequence 3, App1
44	1948.5	63.7	433	4	US-10-456-479-12	Sequence 12, App1
45	1948.5	63.7	433	4	US-10-837-566-3	Sequence 3, App1

ALIGNMENTS

RESULT 1									
US-09-920-954-4									
Sequence 4, Application US/09920954									
Publication No. US20020064854A1									
GENERAL INFORMATION:									
APPLICANT: TAKIWA, MIKIO									
APPLICANT: OKUDA, MITSUYOSHI									
APPLICANT: SAKKI, KATSUHIISA									
APPLICANT: KUBOTA, HIROMI									
APPLICANT: KUBOTA, JUN									
APPLICANT: KAGEYAMA, YASUSHI									
APPLICANT: SHIRAKA, SHITSUM									
APPLICANT: NOMURA, MASAFUMI									
TITLE OF INVENTION: ALKALINE PROTEASE									
FILE REFERENCE: 0327-0832-0PCT									
CURRENT APPLICATION NUMBER: US/09/920,954									
CURRENT FILING DATE: 2001-08-03									
PRIOR APPLICATION NUMBER: 09/509,814									
PRIOR FILING DATE: 2000-04-06									
PRIOR APPLICATION NUMBER: PCT/JP98/04528									
PRIOR FILING DATE: 1998-10-07									
PRIOR APPLICATION NUMBER: JP 9-274570									
PRIOR FILING DATE: 1997-06-08									
NUMBER OF SEQ ID NOS: 24									
SOFTWARE: PatentIn version 3.0									
SEQ ID NO 4									
LENGTH: 639									
TYPE: PRT									
ORGANISM: Bacillus sp.									
US-09-920-954-4									
Query Match									
Best Local Similarity 99.1%; Score 3030; DB 3; Length 639;									
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;									
QY	1	MRKKVPLSVLSAAAILSTV	ALXNPSAGAKRFDDFGIQT	TTXXGFSKQQTGAALF	60				
DB	1	MRKKVPLSVLSAAAILSTV	ALNPNPSAGDARFDLDFG	IQTTTVSGFSKQRTGAALF	60				
QY	61	LVESNVKLKGLKGLSTVP	ANNLGHIXFNGPILBETK	QXLEXTGAKIIDYIDYVI	120				
DB	61	LVESNVKLKGLKGLSTVP	ANNLGHIVQFNGPILBETK	QLEXTGAKIIDYIDYVI	120				
QY	121	VEYEDVYKXXKXIEHVS	VEPYLPHYIIDPLFTK	GSXLYKAAALDTKQNK	VEVLRG 180				
DB	121	VEYEDVYKXXKXIEHVS	VEPYLPHYIIDPLFTK	GSXLYKAAALDTKQNK	VEVLRG 180				
QY	181	IEIXQXXKXNDVXYITAK	PEYKWNDAVARGIVK	ADVAQSSYGLGQGOI	VAADTGLDT 240				
DB	181	IEIXQXXKXNDVXYITAK	PEYKWNDAVARGIVK	ADVAQSSYGLGQGOI	VAADTGLDT 240				

QY 241 GRNDSMEAFRGKLTALYALGRTNANDTNGHTVAGSVLGNKTKGMAPOANLVFQ 300
DB 241 GRNDSMEAFRGKLTALYALGRTNANDTNGHTVAGSVLGNKTKGMAPOANLVFQ 300
QY 301 SIMDSXGGLGGLPSNLQTLFSAQASAGARIHTNSMGAUVNGAYTTDSRVNDYVRKNDMT 360
DB 301 SIMDSXGGLGGLPSNLQTLFSAQASAGARIHTNSMGAUVNGAYTTDSRVNDYVRKNDMT 360
QY 361 ILFAAGNEXPNGGTTISAPGTAKNAITVGATENILRPSFGSYADNINHVAQFSSRGPTKGR 420
DB 361 ILFAAGNEXPNGGTTISAPGTAKNAITVGATENILRPSFGSYADNINHVAQFSSRGPTKGR 420
QY 421 IKPDVMAFGTYILSARSSILAPDSSFWANHDSKYAVMGTSMAPIVAGNVAQLEHFPYKN 480
DB 421 IKPDVMAFGTYILSARSSILAPDSSFWANHDSKYAVMGTSMAPIVAGNVAQLEHFPYKN 480
QY 481 RGITPKPSLLKAAALAGAADXGLGYPNGOGMGRVTLDKSLNVAAYVNESSXLSSTOKATY 540
DB 481 RGITPKPSLLKAAALAGAADXGLGYPNGOGMGRVTLDKSLNVAAYVNESSXLSSTOKATY 540
QY 541 XFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXYVGNDFXXXPXXXNDG 600
DB 541 TFTAATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXYVGNDFXXXPXXXNDG 600
QY 601 RNNVENVFINKPOSGTYYTIEVOAYNVPGPQNFSLATVN 639
DB 601 RNNVENVFINKPOSGTYYTIEVOAYNVPGPQNFSLATVN 639

RESULT 2
US-10-784-870-4

Sequence 4, Application US/10784870
Publication No. US20040142837A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 639
TYPE: PRT
ORGANISM: Bacillus sp.
US-10-784-870-4

Query Match 99.1%; Score 3030; DB 4; Length 639;
Best Local Similarity 93.4%; Pred. No. 4,6e-261;
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 MRKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTITDXGFSKXQGTGAALF 60
DB 1 MRKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTITDXGFSKXQGTGAALF 60
QY 61 LVSESNVTLKGLKGLKLETVPANNTLHIXQFNGPILEETKQKLETKAKILDIYIDYAI 120
DB 61 LVSESNVTLKGLKGLKLETVPANNTLHIXQFNGPILEETKQKLETKAKILDIYIDYAI 120

QY 121 VEYEGDVAXXXXXXIEHVESVREPYLPXYIIDPOLTKGASXLVAXAALDTKQXKEVQLRG 180
DB 121 VEYEGDVAXKVRSLIEHVESVREPYLPKYIIDPOLTKGASLTVKALADTKQXKNEVQLRG 180
QY 181 IEXIAQXXSNDVXYITLAKPEYKVMNDVARGIVRADVQSSYGLYGGQQLVAVADTGLDT 240
DB 181 IEBIAQYASNDVHYITLAKPEYKVMNDVARGIVRADVQSSYGLYGGQQLVAVADTGLDT 240
QY 241 GRNDSMEAFRGKLTALYALGRTNANDTNGHTVAGSVLGNKTKGMAPOANLVFQ 300
DB 241 GRNDSMEAFRGKLTALYALGRTNANDTNGHTVAGSVLGNKTKGMAPOANLVFQ 300
QY 301 SIMDSXGGLGGLPSNLQTLFSAQASAGARIHTNSMGAUVNGAYTTDSRVNDYVRKNDMT 360
DB 301 SIMDSXGGLGGLPSNLQTLFSAQASAGARIHTNSMGAUVNGAYTTDSRVNDYVRKNDMT 360
QY 361 ILFAAGNEXPNGGTTISAPGTAKNAITVGATENILRPSFGSYADNINHVAQFSSRGPTKGR 420
DB 361 ILFAAGNEXPNGGTTISAPGTAKNAITVGATENILRPSFGSYADNINHVAQFSSRGPTKGR 420
QY 421 IKPDVMAFGTYILSARSSILAPDSSFWANHDSKYAVMGTSMAPIVAGNVAQLEHFPYKN 480
DB 421 IKPDVMAFGTYILSARSSILAPDSSFWANHDSKYAVMGTSMAPIVAGNVAQLEHFPYKN 480
QY 481 RGITPKPSLLKAAALAGAADXGLGYPNGOGMGRVTLDKSLNVAAYVNESSXLSSTOKATY 540
DB 481 RGITPKPSLLKAAALAGAADXGLGYPNGOGMGRVTLDKSLNVAAYVNESSXLSSTOKATY 540
QY 541 XFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXYVGNDFXXXPXXXNDG 600
DB 541 TFTAATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXYVGNDFXXXPXXXNDG 600
QY 601 RNNVENVFINKPOSGTYYTIEVOAYNVPGPQNFSLATVN 639
DB 601 RNNVENVFINKPOSGTYYTIEVOAYNVPGPQNFSLATVN 639

RESULT 3

US-09-920-954-6
Sequence 6, Application US/09920954
Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-920-954-6

Query Match 98.8%; Score 3022; DB 3; Length 640;
Best Local Similarity 93.3%; Pred. No. 2,4e-260;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 2 RKKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTITDXGFSKXQGTGAALF 61
DB 2 RKKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTITDXGFSKXQGTGAALF 61

DB 3 KKKKVFSLVLSAAAIISTVALSNPSAGARNPDLDFKGIQTTTDAKGFSGOGTGAAPL 62
QY 62 VSEENVKLMKKKKLLETPANNKHLHXOFGPILEETKQXLEXTAKLIDYIPDYAYIV 121
DB 63 VSEENVKLPKQKLETPANNKHLHXOFGPILEETKQXLEXTAKLIDYIPDYAYIV 122
QY 122 EYEGDVYXXXXIEHVESVEPYLPYXIDPOLFTKASLVKAXALDTQXNKEVOLRG 181
DB 123 EYEGDVYKATSTIEHVESVEPYLPYRIDPOLFTKASLVKAXALDTQXNKEVOLRG 182
QY 182 EXIAQKXSNVYXYITAKEBYKMDVAGIYKADVAQSSYGLYGGQIYVAVADTGLDTG 241
DB 183 EYIAQFALISNDVLYITAKEBYKMDVAGIYKADVAQSSYGLYGGQIYVAVADTGLDTG 242
QY 242 RNDSSMHEAFRGKITLVALGRTNNANDTNGHGTVAAGSVLGNKXNKMAPOANLVFOS 301
DB 243 RNDSSMHEAFRGKITLVALGRTNNANDTNGHGTVAAGSVLGNKXNKMAPOANLVFOS 302
QY 302 IMDSXGGLGGLPSNLOTLFSAQXASAGARIHTNSWGAANVGAAYTTDSRNDVYRKNDMTI 361
DB 303 IMDSXGGLGGLPSNLOTLFSAQXASAGARIHTNSWGAANVGAAYTTDSRNDVYRKNDMTI 362
QY 362 LPAAGNEXENGCTISAPGTAKNAITVGATENLRPSFGSYADININHAOFSSRGPYDGR 421
DB 363 LPAAGNEXENGCTISAPGTAKNAITVGATENLRPSFGSYADININHAOFSSRGPYDGR 422
QY 422 KPDVAPGTXYLLSASSSLAPDSSFWANHDSKAYWCGTSMATPIVAGNVAQLEHVPK 481
DB 423 KPDVAPGTXYLLSASSSLAPDSSFWANHDSKAYWCGTSMATPIVAGNVAQLEHVPK 482
QY 482 GTPPKSLKALILAGAADXGGLYENGNGMGVTLDKSLANVYVNESSSLSTSOKATY 541
DB 483 GTPPKSLKALILAGAADXGGLYENGNGMGVTLDKSLANVYVNESSSLSTSOKATY 542
QY 542 FTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXVYVGNDFXPXXNMDGR 601
DB 543 FTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXVYVGNDFXPXXNMDGR 602
QY 602 NNVENVFINKPOSGTYTIEVOAYNVPGPQTSLSIAVN 639
DB 603 NNVENVFINKPOSGTYTIEVOAYNVPGPQTSLSIAVN 640
RESULT 4
US-10-456-479-4
Sequence 4, Application US/10456479
Publication No. US20040072321A1
GENERAL INFORMATION:
APPLICANT: SATO, TSUYOSHI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: TAKIMURA, YASUSHI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: NOMURA, MASAFUMI
APPLICANT: KOBAYASHI, TOHRU
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 238700USO
CURRENT APPLICATION NUMBER: US/10/456,479
CURRENT FILING DATE: 2003-06-09
PRIOR APPLICATION NUMBER: JP 2002-186387
PRIOR FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: JP 2002-304232
PRIOR FILING DATE: 2002-10-18
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp. KSM-KP43
US-10-456-479-4
Query Match 98.8%; Score 3022; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 2.4e-260;

*Common inventory
with my appl.
However there is no
assigned for my appl.
Jan 11
5/25/06*

Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
QY 2 RKKKVFSLVLSAAAIISTVALSNPSAGARNPDLDFKGIQTTTDXKGFSGOGTGAAPL 61
DB 3 KKKKVFSLVLSAAAIISTVALSNPSAGARNPDLDFKGIQTTTDAKGFSGOGTGAAPL 62
QY 62 VSEENVKLMKKKKLLETPANNKHLHXOFGPILEETKQXLEXTAKLIDYIPDYAYIV 121
DB 63 VSEENVKLPKQKLETPANNKHLHXOFGPILEETKQXLEXTAKLIDYIPDYAYIV 122
QY 122 EYEGDVYXXXXIEHVESVEPYLPYXIDPOLFTKASLVKAXALDTQXNKEVOLRG 181
DB 123 EYEGDVYKATSTIEHVESVEPYLPYRIDPOLFTKASLVKAXALDTQXNKEVOLRG 182
QY 182 EXIAQKXSNVYXYITAKEBYKMDVAGIYKADVAQSSYGLYGGQIYVAVADTGLDTG 241
DB 183 EYIAQFALISNDVLYITAKEBYKMDVAGIYKADVAQSSYGLYGGQIYVAVADTGLDTG 242
QY 242 RNDSSMHEAFRGKITLVALGRTNNANDTNGHGTVAAGSVLGNKXNKMAPOANLVFOS 301
DB 243 RNDSSMHEAFRGKITLVALGRTNNANDTNGHGTVAAGSVLGNKXNKMAPOANLVFOS 302
QY 302 IMDSXGGLGGLPSNLOTLFSAQXASAGARIHTNSWGAANVGAAYTTDSRNDVYRKNDMTI 361
DB 303 IMDSXGGLGGLPSNLOTLFSAQXASAGARIHTNSWGAANVGAAYTTDSRNDVYRKNDMTI 362
QY 362 LPAAGNEXENGCTISAPGTAKNAITVGATENLRPSFGSYADININHAOFSSRGPYDGR 421
DB 363 LPAAGNEXENGCTISAPGTAKNAITVGATENLRPSFGSYADININHAOFSSRGPYDGR 422
QY 422 KPDVAPGTXYLLSASSSLAPDSSFWANHDSKAYWCGTSMATPIVAGNVAQLEHVPK 481
DB 423 KPDVAPGTXYLLSASSSLAPDSSFWANHDSKAYWCGTSMATPIVAGNVAQLEHVPK 482
QY 482 GTPPKSLKALILAGAADXGGLYENGNGMGVTLDKSLANVYVNESSSLSTSOKATY 541
DB 483 GTPPKSLKALILAGAADXGGLYENGNGMGVTLDKSLANVYVNESSSLSTSOKATY 542
QY 542 FTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXVYVGNDFXPXXNMDGR 601
DB 543 FTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXVYVGNDFXPXXNMDGR 602
QY 602 NNVENVFINKPOSGTYTIEVOAYNVPGPQTSLSIAVN 639
DB 603 NNVENVFINKPOSGTYTIEVOAYNVPGPQTSLSIAVN 640
RESULT 5
US-10-784-870-6
Sequence 6, Application US/10784870
Publication No. US20040142837A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-OPCT
CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/J998/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6

LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-10-784-870-6

Query Match 98.8%; Score 3022; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 2,4e-260;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```
QY 2 RKKKVFLSVLSAAALISTVALXNPSAGARXFDLDFKGIQTTDXGFSKXQGTGAAPL 61
DB 3 KKKKVFSLVLSAAALISTVALSNPSAGARFDFLDFKGIQTTDXGFSKXQGTGAAPL 62
QY 62 VESENVKLKXKLKLTVPANNKLIHQFNGLILEETKQLETKAKILDYIPDYATIV 121
DB 63 VESENVKLKPKLQKLETVANNKLIHQFNGLILEETKQLETKAKILDYIPDYATIV 122
QY 122 EYEGDVXSKXXXIHEVESVEPYLPYXIIDPOLFTKGASXLVAKAXALDTKXNKEVOLRGI 181
DB 123 EYEGDVKSATSTIEHVESVEPYLPYXIIDPOLFTKGASXLVAKAXALDTKXNKEVOLRGI 182
QY 182 EXIAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAVADTGLDTG 241
DB 183 EQIAQFALSNDVLYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAVADTGLDTG 242
QY 242 RNDSSMHEAFRGKLTALYALGRTNNANDTNGHGHVAGSVLGNGSTKGMAPQANLVFQS 301
DB 243 RNDSSMHEAFRGKLTALYALGRTNNANDTNGHGHVAGSVLGNGSTKGMAPQANLVFQS 302
QY 302 IMDSXGGLGGLPSNLQTLFSGQXSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 361
DB 303 IMDSGGLGGLPSNLQTLFSGQXSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 362
QY 362 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR 421
DB 363 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR 422
QY 422 KPDVMAFGTXILSARSSLAPDSSFWANHDSKYAYMGTSMAPIVAGNVADLRHFPYNR 481
DB 423 KPDVMAFGTXILSARSSLAPDSSFWANHDSKYAYMGTSMAPIVAGNVADLRHFPYNR 482
QY 482 GTPKPSLLKXALLAGADGIGYPNGOGMGRVTLDSLVNAVYNESSXLSSTOKATYX 541
DB 483 GTPKPSLLKXALLAGADGIGYPNGOGMGRVTLDSLVNAVYNESSXLSSTOKATYX 542
QY 542 FTATAGKPLKISLVWSDAPASTTASVTLVNDLIVITAPNGTYVGNDFXXPKXXNMDGR 601
DB 543 FTATAGKPLKISLVWSDAPASTTASVTLVNDLIVITAPNGTYVGNDFPSYNDNMDGR 602
QY 602 NNVENVFINKPQSGTYTIEVOAYNVVPQPFSLAIYN 639
DB 603 NNVENVFINKPQSGTYTIEVOAYNVVPQPFSLAIYN 640
```

RESULT 6
US-10-820-712A-3
Sequence 3, Application US/10820712A
Publication No. US20050026804A1

GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuoshi
APPLICANT: Izawa, Yoshihumi
APPLICANT: Kobayashi, Toku
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-USO
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106708
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patent version 3.2

Search
Result 6

SEQ ID NO 3
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp. KSM-KP43
US-10-820-712A-3

Query Match 98.8%; Score 3022; DB 5; Length 640;
Best Local Similarity 93.3%; Pred. No. 2,4e-260;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```
QY 2 RKKKVFLSVLSAAALISTVALXNPSAGARXFDLDFKGIQTTDXGFSKXQGTGAAPL 61
DB 3 KKKKVFSLVLSAAALISTVALSNPSAGARFDFLDFKGIQTTDXGFSKXQGTGAAPL 62
QY 62 VESENVKLKXKLKLTVPANNKLIHQFNGLILEETKQLETKAKILDYIPDYATIV 121
DB 63 VESENVKLKPKLQKLETVANNKLIHQFNGLILEETKQLETKAKILDYIPDYATIV 122
QY 122 EYEGDVXSKXXXIHEVESVEPYLPYXIIDPOLFTKGASXLVAKAXALDTKXNKEVOLRGI 181
DB 123 EYEGDVKSATSTIEHVESVEPYLPYXIIDPOLFTKGASXLVAKAXALDTKXNKEVOLRGI 182
QY 182 EXIAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAVADTGLDTG 241
DB 183 EQIAQFALSNDVLYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAVADTGLDTG 242
QY 242 RNDSSMHEAFRGKLTALYALGRTNNANDTNGHGHVAGSVLGNGSTKGMAPQANLVFQS 301
DB 243 RNDSSMHEAFRGKLTALYALGRTNNANDTNGHGHVAGSVLGNGSTKGMAPQANLVFQS 302
QY 302 IMDSXGGLGGLPSNLQTLFSGQXSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 361
DB 303 IMDSGGLGGLPSNLQTLFSGQXSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 362
QY 362 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR 421
DB 363 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR 422
QY 422 KPDVMAFGTXILSARSSLAPDSSFWANHDSKYAYMGTSMAPIVAGNVADLRHFPYNR 481
DB 423 KPDVMAFGTXILSARSSLAPDSSFWANHDSKYAYMGTSMAPIVAGNVADLRHFPYNR 482
QY 482 GTPKPSLLKXALLAGADGIGYPNGOGMGRVTLDSLVNAVYNESSXLSSTOKATYX 541
DB 483 GTPKPSLLKXALLAGADGIGYPNGOGMGRVTLDSLVNAVYNESSXLSSTOKATYX 542
QY 542 FTATAGKPLKISLVWSDAPASTTASVTLVNDLIVITAPNGTYVGNDFXXPKXXNMDGR 601
DB 543 FTATAGKPLKISLVWSDAPASTTASVTLVNDLIVITAPNGTYVGNDFPSYNDNMDGR 602
QY 602 NNVENVFINKPQSGTYTIEVOAYNVVPQPFSLAIYN 639
DB 603 NNVENVFINKPQSGTYTIEVOAYNVVPQPFSLAIYN 640
```

RESULT 7

US-10-820-714A-3
Sequence 3, Application US/10820714A
Publication No. US20050214922A1

GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuoshi
APPLICANT: Kobayashi, Toku
APPLICANT: Sumitomo, Nobuyuki
APPLICANT: Takimura, Yasushi
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251697USO
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106709
NUMBER OF SEQ ID NOS: 24

APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-10-784-870-8

Query Match 98.8%; Score 3021; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 2,9e-260;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 2 RKKKVFSLVSAALISTVVALKXNPAGKARXFDLPKGIQTITDXKGFSGKQQTGAAPL 61
DB 3 KKKKVFSLVSAALISTVVALKXNPAGKARXFDLPKGIQTITDXKGFSGKQQTGAAPL 62
QY 62 VESENVKLLKGLKLETPVANNKHIHQNGPILLETQKLEXTGAKLIDYIPDYAYIV 121
DB 63 VESENVKLLKGLKLETPVANNKHIHQNGPILLETQKLEXTGAKLIDYIPDYAYIV 122
QY 122 EYEGDVKXXXKIEHVESVEPYLPYXXIDPOLFTKGASXLVAKALDTKXNKEVOLRGI 181
DB 123 EYEGDVKXSATSTIEHVESVEPYLPYRIDPOLFTKGASXLVAKALDTKXNKEVOLRGI 182
QY 182 EXIAQXXXNDXYITAKPEYKVMNDVARGIVKADVAOSSVGLYQGGQIVNADTGLDNG 241
DB 183 EYEGDVKXSATSTIEHVESVEPYLPYRIDPOLFTKGASXLVAKALDTKXNKEVOLRGI 242
QY 242 RNDSSMHEAFRGKITALVALGRTNNANDTNGHGTAVAGSVLNGXTNKGMAPOANLVPQS 301
DB 243 RNDSSMHEAFRGKITALVALGRTNNANDTNGHGTAVAGSVLNGXTNKGMAPOANLVPQS 302
QY 302 IMDSYGGJGGLPSNTQTLFSGAKXSGARITHNSMGAAVNGAYTTDSRVNDYVRKNDMTI 361
DB 303 IMDSYGGJGGLPSNTQTLFSGAKXSGARITHNSMGAAVNGAYTTDSRVNDYVRKNDMTI 362
QY 362 LPAAGNEXPNGGTISAPGAKKAIIVGATENLRPSFGVADININVAOPSSRGPTKDGRI 421
DB 363 LPAAGNEXPNGGTISAPGAKKAIIVGATENLRPSFGVADININVAOPSSRGPTKDGRI 422
QY 422 KPDVVAPEGTILISASSSLAPDSSFWANHDSKAYVNGTSMATPIYAGVNAQLREHFVNQR 481
DB 423 KPDVVAPEGTILISASSSLAPDSSFWANHDSKAYVNGTSMATPIYAGVNAQLREHFVNQR 482
QY 482 GITPEPSLLKALIGAADKGLGYNGNGQWGRVTLDSLVNAVYNESXLSSTQKATYX 541
DB 483 GITPEPSLLKALIGAADKGLGYNGNGQWGRVTLDSLVNAVYNESXLSSTQKATYX 542
QY 542 FTATGKPLKTSLVNSDAPASTTASVTLVNDLVTITPNGXYVNGDPKXPKXNNMGR 601
DB 543 FTATGKPLKTSLVNSDAPASTTASVTLVNDLVTITPNGXYVNGDPKXPKXNNMGR 602
QY 602 NNVENVPINXPOSGTYTIEVQAYNPVGPQXFSIAIVN 639
DB 603 NNVENVPINXPOSGTYTIEVQAYNPVGPQXFSIAIVN 640

RESULT 10
US-09-920-954-1
Sequence 1, Application US/09920954

Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUMISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature

```
; LOCATION: (133)..(133)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (146)..(146)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (148)..(148)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (160)..(160)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (165)..(165)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (172)..(172)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (183)..(183)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (187)..(187)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (188)..(188)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (189)..(189)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (194)..(194)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (286)..(286)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (306)..(306)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (324)..(324)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (369)..(369)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (431)..(431)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (501)..(501)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (531)..(531)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (541)..(541)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (584)..(584)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (591)..(591)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (592)..(592)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (594)..(594)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (595)..(595)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (596)..(596)
```

```
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (611)..(611)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: (632)..(632)
; OTHER INFORMATION: Xaa is any amino acid
US-09-920-954-1

Query Match      98.6%; Score 3017; DB 3; Length 639;
Best Local Similarity 100.0%; Pred. No. 6,6e-260;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKKKVFLSVLSAAAILSTVALNPSAGAKKYPDLDFKGIQTTHXKXGFSKXQGTGAAP 60
DB      1 MKKKVFLSVLSAAAILSTVALNPSAGAKKYPDLDFKGIQTTHXKXGFSKXQGTGAAP 60
QY      61 LVSENVVQLXKGLXKKLETPVANNKLIHQFNGPILEETKXKLEXTGAKILDIYIPYAYI 120
DB      61 LVSENVVQLXKGLXKKLETPVANNKLIHQFNGPILEETKXKLEXTGAKILDIYIPYAYI 120
QY      121 VEYEGDVASXXXXIHVESVEPYLPYXIDPOLFTKGASXLVKAALDTKXNKVEQLNG 180
DB      121 VEYEGDVASXXXXIHVESVEPYLPYXIDPOLFTKGASXLVKAALDTKXNKVEQLNG 180
QY      181 IEXIAQXXXSNDVXYITAKPEYKVNADVAGIVKADVAQSSYGLYGQGIIVAVDTGLDT 240
DB      181 IEXIAQXXXSNDVXYITAKPEYKVNADVAGIVKADVAQSSYGLYGQGIIVAVDTGLDT 240
QY      241 GRNDSMHEAFRGKITALYALGRTNANDTNHGHVAGSVGNKXTKGMAPQANTVPO 300
DB      241 GRNDSMHEAFRGKITALYALGRTNANDTNHGHVAGSVGNKXTKGMAPQANTVPO 300
QY      301 SIMDSXGGLGGLPSNLQTLFSQXSAGARIHTNSWGAANVGAYTTDSRVDDYRKNDMT 360
DB      301 SIMDSXGGLGGLPSNLQTLFSQXSAGARIHTNSWGAANVGAYTTDSRVDDYRKNDMT 360
QY      361 ILFAAGNEXPNGGTSAPGTAKNAITVGATENLRPSFSYADNINHVAFQSSRGPTKGR 420
DB      361 ILFAAGNEXPNGGTSAPGTAKNAITVGATENLRPSFSYADNINHVAFQSSRGPTKGR 420
QY      421 IKPDVMAFGTILSRSSSLAPDSSFWANHDSRYATMGTSMAPIVAGVADLRHPVYN 480
DB      421 IKPDVMAFGTILSRSSSLAPDSSFWANHDSRYATMGTSMAPIVAGVADLRHPVYN 480
QY      481 RGITPKPSILKXALJAGAADXGLGYPNGNGGRTYTLDSLVNAVYNESSXLSSTOKATY 540
DB      481 RGITPKPSILKXALJAGAADXGLGYPNGNGGRTYTLDSLVNAVYNESSXLSSTOKATY 540
QY      541 XFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXVYVNDPXXPXXNDMG 600
DB      541 XFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXVYVNDPXXPXXNDMG 600
QY      601 RNNVENVPINXPOSGTITIEVQAYNVPGPQXFSALAYN 639
DB      601 RNNVENVPINXPOSGTITIEVQAYNVPGPQXFSALAYN 639

RESULT 11
US-10-784-870-1
; Sequence 1, Application US/10784870
; Publication No. US20040142837A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGIYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
```

```

CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPE: PR1
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (126)..(126)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid

```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-10-784-870-1
```

Query Match 98.6%; Score 3017; DB 4; Length 639;
Best Local Similarity 100.0%; Pred. No. 6.6e-260; Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRKKVPLSVLSAAIILSTVALXNPSAGARXFDLDFKGIOTTDXKXGFSKXONGAAAF 60
DB 1 MRKKVPLSVLSAAIILSTVALXNPSAGARXFDLDFKGIOTTDXKXGFSKXONGAAAF 60
QY 61 LVSEENVQLXKGLXKGLXKGLXKGLXKGLXKGLXKGLXKGLXKGLXKGLXKGLXKGLX 120
DB 61 LVSEENVQLXKGLXKGLXKGLXKGLXKGLXKGLXKGLXKGLXKGLXKGLXKGLXKGLX 120
QY 121 VVEGDVVSXXXXIEHVESVEPYLPYXIDPOLFTKGASXLVXAXALDTKXNKEVOLRG 180
DB 121 VVEGDVVSXXXXIEHVESVEPYLPYXIDPOLFTKGASXLVXAXALDTKXNKEVOLRG 180
QY 181 IEXIOXXXXNDVYITAKPEYKMNNDVARGIVKADVQSSGYLGQGGIIVAVADTGLDT 240
DB 181 IEXIOXXXXNDVYITAKPEYKMNNDVARGIVKADVQSSGYLGQGGIIVAVADTGLDT 240
QY 241 GRNDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVGNGXTNKGMAPOANLVFQ 300
DB 241 GRNDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVGNGXTNKGMAPOANLVFQ 300
QY 301 SIMDSXGIGIGIPSNLQTFPSQAXSAGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDT 360
DB 301 SIMDSXGIGIGIPSNLQTFPSQAXSAGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDT 360
QY 361 ILFAAGNEXPNGGTISAPGTAKNAITVGTENLRPSFGSYANINHVQPFSSRGPTKOR 420
DB 361 ILFAAGNEXPNGGTISAPGTAKNAITVGTENLRPSFGSYANINHVQPFSSRGPTKOR 420
QY 421 IKPDVMAPOTXILSARSSSLAPDSFWANHDSKYAVWGTSMAITPIVAGVNAQLREHFVN 480
DB 421 IKPDVMAPOTXILSARSSSLAPDSFWANHDSKYAVWGTSMAITPIVAGVNAQLREHFVN 480
```

```
QY 481 RGITPSPILKALJAGADKGLGYPNNGQGRVTLDDKSLVAVYNESSXLTSGKATY 540
DB 481 RGITPSPILKALJAGADKGLGYPNNGQGRVTLDDKSLVAVYNESSXLTSGKATY 540
QY 541 XFTATGKPLKSLVWSDAPASTTASVTLVNDLDTVITAPNGTXYVGNDFXXPKXNMG 600
DB 541 XFTATGKPLKSLVWSDAPASTTASVTLVNDLDTVITAPNGTXYVGNDFXXPKXNMG 600
QY 601 RNNVENFINXPOSGTITIEVQAYNVPGQXFSIAVN 639
DB 601 RNNVENFINXPOSGTITIEVQAYNVPGQXFSIAVN 639
```

```
RESULT 12
US-09-920-954-2
Sequence 2, Application US/09920954
Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKI, KATSUHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PR
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (90)..(90)
```

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid

NAME/KEY: misc feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-920-954-2

Query Match 98.3%, Score 3007, DB 3, length 640;
Best Local Similarity 100.0%; Pred. No. 5, 2e-259;
Matches 637, Conservative 0, Mismatches 0, Indels 0, Gaps 0,

QY 3 KKKVFLSVLSAAALISVALXNPAGAXRPFDDFKGIQTITDXXGFSKQXQTGAAPLV 62
DB 4 KKKVFLSVLSAAALISVALXNPAGAXRPFDDFKGIQTITDXXGFSKQXQTGAAPLV 63
QY 63 ESENVKXKXGIXKXLEVPANRKLHIQPNPILEETKQXLETKGAILDVIPIYATYVE 122
DB 64 ESENVKXKXGIXKXLEVPANRKLHIQPNPILEETKQXLETKGAILDVIPIYATYVE 123
QY 123 YEGDVXSKXXXIHEVESVEPYLPYXXIDPOLFTFGASXLVAXALDTRQXNKEVQLNGIE 182
DB 124 YEGDVXSKXXXIHEVESVEPYLPYXXIDPOLFTFGASXLVAXALDTRQXNKEVQLNGIE 183
QY 183 XIAQXXXSNDVXYITTAPEYKVMNDVARGIKADVAQSSYGLYGQGIIVAAVDTGLDTGR 242
DB 184 XIAQXXXSNDVXYITTAPEYKVMNDVARGIKADVAQSSYGLYGQGIIVAAVDTGLDTGR 243
QY 243 NDSMHEAFRKTITVALGRTNNANDTNGHTVAGSYLGNKTNMGMAPOALVPOSI 302
DB 244 NDSMHEAFRKTITVALGRTNNANDTNGHTVAGSYLGNKTNMGMAPOALVPOSI 303
QY 303 MDSXGGLGGLPSNQTLPQXKSAGARIHTNSMGAANVGAATTSRNVDDVVRKNDTIL 362
DB 304 MDSXGGLGGLPSNQTLPQXKSAGARIHTNSMGAANVGAATTSRNVDDVVRKNDTIL 363
QY 363 FAAGNEXPNGGTISAPGTAKNAITVGATEMLRPSFGSYADININVAQFSSRGPTKGRIK 422
DB 364 FAAGNEXPNGGTISAPGTAKNAITVGATEMLRPSFGSYADININVAQFSSRGPTKGRIK 423
QY 423 PDVNAFGTXILISASSSLAPDSSFPANHDSKAYVNGTSMATPIYAGVNAQLREHFNVRG 482
DB 424 PDVNAFGTXILISASSSLAPDSSFPANHDSKAYVNGTSMATPIYAGVNAQLREHFNVRG 483
QY 483 ITPKPSLLKALIGAADXGLGYPNGNGWRVTLDSLVNAVAYVNESSXLSSTQKATYXF 542
DB 484 ITPKPSLLKALIGAADXGLGYPNGNGWRVTLDSLVNAVAYVNESSXLSSTQKATYXF 543
QY 543 TATAGKPLKISLVMSDAPASTTASVTLVNDLIVITAPNGTXVYVGNDFXXPXKXNWDGRN 602
DB 544 TATAGKPLKISLVMSDAPASTTASVTLVNDLIVITAPNGTXVYVGNDFXXPXKXNWDGRN 603

QY 603 NVNVPINXPOSTGTITIEVOANVPVGPQXFSALAYN 639
Db 604 NVNVPINXPOSTGTITIEVOANVPVGPQXFSALAYN 640

RESULT 13
US-10-784-870-2
Sequence 2. Application US/10784870
Publication No. US20040142837A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid

OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid

FEATURE:
NAME/KEY: misc_feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-10-784-870-2

Query Match 98.3%; Score 3007; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 5.2e-259; Indels 0; Gaps 0;
Matches 637; Conservative 0; Mismatches 0;

Qy 3 KKKVSLVLSAAAILSTVALXNPSAGXARXPDLDKGIQTTTDXGFGSKQGTGAALV 62
Db 4 KKKVSLVLSAAAILSTVALXNPSAGXARXPDLDKGIQTTTDXGFGSKQGTGAALV 63
Qy 63 ESENKXKXGKXKXKXETVPANNKCHIXQFNGPILEETXOXILEXTAKLIDYIPDYAYIVE 122
Db 64 ESENKXKXGKXKXKXETVPANNKCHIXQFNGPILEETXOXILEXTAKLIDYIPDYAYIVE 123

Qy 123 YEGDVXSKXXKIIEHVESVEPLPYXXIDPOLFTKGASLYVQXALDTRKXNKKEVQLRGIE 182
Db 124 YEGDVXSKXXKIIEHVESVEPLPYXXIDPOLFTKGASLYVQXALDTRKXNKKEVQLRGIE 183
Qy 183 XIAOXXSNDVXYITAKPEYKAMDVARGIYKADVAQSSYGLYGGQGVAAVADTGLDGR 242
Db 184 XIAOXXSNDVXYITAKPEYKAMDVARGIYKADVAQSSYGLYGGQGVAAVADTGLDGR 243
Qy 243 NDSMHEAFRGKITALVALGRTNANDTNGHGTIVAGSVLGNKXTNKGMAPOANLVFQSI 302
Db 244 NDSMHEAFRGKITALVALGRTNANDTNGHGTIVAGSVLGNKXTNKGMAPOANLVFQSI 303
Qy 303 MDSXGGLGLPSNQLTFESQXKSAGARHTMSKGAANVGATTTBSRNVDDVVRKNDITL 362
Db 304 MDSXGGLGLPSNQLTFESQXKSAGARHTMSKGAANVGATTTBSRNVDDVVRKNDITL 363
Qy 363 FAAGNEXPNGGTISAPGTAKNAITVGATENTLRFSGSYADNINHVAQPSRGPYKGRIRK 422
Db 364 FAAGNEXPNGGTISAPGTAKNAITVGATENTLRFSGSYADNINHVAQPSRGPYKGRIRK 423
Qy 423 PDVNAFGTXIILSABSSLPADSSFWANHDSKYAYNGGTSMATPIVAGNVAQLREHFVKNRG 482
Db 424 PDVNAFGTXIILSABSSLPADSSFWANHDSKYAYNGGTSMATPIVAGNVAQLREHFVKNRG 483
Qy 483 ITPKPSILKALILGAADXGYPNGNGKWRVTLDLSINAYYNESKXISTSQATYXF 542
Db 484 ITPKPSILKALILGAADXGYPNGNGKWRVTLDLSINAYYNESKXISTSQATYXF 543
Qy 543 TATAGKPLKISLVNSDAPASTTASVTLVNDLDTLITAPNGTXYYGNDPFXKPKXXNMDGRN 602
Db 544 TATAGKPLKISLVNSDAPASTTASVTLVNDLDTLITAPNGTXYYGNDPFXKPKXXNMDGRN 603
Qy 603 NVENVFINKPQSGTYTTEVQAINVPVGPQXFSALIVN 639
Db 604 NVENVFINKPQSGTYTTEVQAINVPVGPQXFSALIVN 640

RESULT 14
US-09-985-689A-1
; Sequence 1, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAMA, AKINORI
; APPLICANT: KAGIYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKETI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-1

Query Match 70.4%; Score 2155; DB 3; Length 434;
Best Local Similarity 96.3%; Pred. No. 2.9e-183; Indels 0; Gaps 0;
Matches 418; Conservative 0; Mismatches 16;

Qy 206 NDVARGIVKADVAQSSYGLYGGQGVAAVADTGLDGRNDSMHEAFRGKITALVALGRTN 265

Db 1 NDVARGIVADVAQSSYGLYGCGQIVAVADTGLDTRGRNDSMHBAFRGKITLALYALGRTN 60
Qy 266 NADNTGHTGVAGSVTLGNKXTTKGMAPOANLVPQSIIMSGXGGLPSNLQTLFEOAXS 325
Db 61 NADNTGHTGVAGSVTLGNKXTTKGMAPOANLVPQSIIMSGXGGLPSNLQTLFEOAXS 120
Qy 326 AGARHTNSWGAANVGAATTTDSRNVDDYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
Db 121 AGARHTNSWGAANVGAATTTDSRNVDDYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
Qy 386 TVGATENLPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPTYLISARSSLAPDSSP 445
Db 181 TVGATENLPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPTYLISARSSLAPDSSP 240
Qy 446 WANHDSKYAYMGTSMAATPIVAGNVAQLREHFVKRNGITPKPSLLKALIIAGAADVGLGY 505
Db 241 WANHDSKYAYMGTSMAATPIVAGNVAQLREHFVKRNGITPKPSLLKALIIAGAADVGLGY 300
Qy 506 PNGNQGMRVTLDDKSLNVAIYVNESSXLSTSQKATYXFTATAGKPKLISLWSDAPASTTA 565
Db 301 PNGNQGMRVTLDDKSLNVAIYVNESSXLSTSQKATYXFTATAGKPKLISLWSDAPASTTA 360
Qy 566 SVTLVNDLVLVTAPNGITYVGNDFPKXPKXNMWDGRNNVENFINSPOGTYTIEVOAYN 625
Db 361 SVTLVNDLVLVTAPNGITYVGNDFPKXPKXNMWDGRNNVENFINSPOGTYTIEVOAYN 420
Qy 626 VEVGPQXFSLATVN 639
Db 421 VEVGPQXFSLATVN 434

RESULT 15
US-09-985-689A-2

/ Sequence 2, Application US/09985689A
/ Publication No. US20030022351A1
/ GENERAL INFORMATION:
/ APPLICANT: HATADA, YUJI
/ APPLICANT: OGAWA, AKINORI
/ APPLICANT: KAGEYAMA, YASUSHI
/ APPLICANT: SATO, TSUYOSHI
/ APPLICANT: ARAKI, HIROYUKI
/ APPLICANT: SUMITOMO, NOBUYUKI
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SAKKI, KATSUHIISA
/ TITLE OF INVENTION: Alkaline proteases
/ FILE REFERENCE: 215483US0
/ CURRENT APPLICATION NUMBER: US/09/985,689A
/ CURRENT FILING DATE: 2002-07-01
/ PRIOR APPLICATION NUMBER: JP P2000-355166
/ PRIOR FILING DATE: 2000-11-22
/ PRIOR APPLICATION NUMBER: JP P2001-114048
/ PRIOR FILING DATE: 2001-04-12
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 434
/ TYPE: PRT
/ ORGANISM: Bacillus sp.
US-09-985-689A-2

Query Match 70.4%; Score 2155; DB 3; Length 434;
Best Local Similarity 96.3%; Pred. No. 2.9e-183;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 206 NDVARGIVADVAQSSYGLYGCGQIVAVADTGLDTRGRNDSMHBAFRGKITLALYALGRTN 265
Db 1 NDVARGIVADVAQSSYGLYGCGQIVAVADTGLDTRGRNDSMHBAFRGKITLALYALGRTN 60
Qy 266 NADNTGHTGVAGSVTLGNKXTTKGMAPOANLVPQSIIMSGXGGLPSNLQTLFEOAXS 325
Db 61 NADNTGHTGVAGSVTLGNKXTTKGMAPOANLVPQSIIMSGXGGLPSNLQTLFEOAXS 120
Qy 326 AGARHTNSWGAANVGAATTTDSRNVDDYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385

Db 121 AGARHTNSWGAANVGAATTTDSRNVDDYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
Qy 386 TVGATENLPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPTYLISARSSLAPDSSP 445
Db 181 TVGATENLPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPTYLISARSSLAPDSSP 240
Qy 446 WANHDSKYAYMGTSMAATPIVAGNVAQLREHFVKRNGITPKPSLLKALIIAGAADVGLGY 505
Db 241 WANHDSKYAYMGTSMAATPIVAGNVAQLREHFVKRNGITPKPSLLKALIIAGAADVGLGY 300
Qy 506 PNGNQGMRVTLDDKSLNVAIYVNESSXLSTSQKATYXFTATAGKPKLISLWSDAPASTTA 565
Db 301 PNGNQGMRVTLDDKSLNVAIYVNESSXLSTSQKATYXFTATAGKPKLISLWSDAPASTTA 360
Qy 566 SVTLVNDLVLVTAPNGITYVGNDFPKXPKXNMWDGRNNVENFINSPOGTYTIEVOAYN 625
Db 361 SVTLVNDLVLVTAPNGITYVGNDFPKXPKXNMWDGRNNVENFINSPOGTYTIEVOAYN 420
Qy 626 VEVGPQXFSLATVN 639
Db 421 VEVGPQXFSLATVN 434

Search completed: April 7, 2006, 18:03:40
Job time : 120.407 secs

This Page Blank (uspto)

November 2005

Published_Applications_Nucleic Acid and Published_Applications_Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 17:59:52 ; Search time 16.4871 Seconds
(without alignments)
1208.926 Million cell updates/sec

Title: US-10-784-870-1

Perfect score: 3059

Sequence: 1 MRKKVPLSVLSAAILSTV.....EYQAVNPVGPQXFLATVN 639

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 184161 seqs, 31191982 residues

Total number of hits satisfying chosen parameters: 184161

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA New:*
1: /SIDS5/ptodata/1/pubppa/US08_NEW_PUB.pap:*
2: /SIDS5/ptodata/1/pubppa/US06_NEW_PUB.pap:*
3: /SIDS5/ptodata/1/pubppa/US07_NEW_PUB.pap:*
4: /SIDS5/ptodata/1/pubppa/PC7_NEW_PUB.pap:*
5: /SIDS5/ptodata/1/pubppa/US09_NEW_PUB.pap:*
6: /SIDS5/ptodata/1/pubppa/US10_NEW_PUB.pap:*
7: /SIDS5/ptodata/1/pubppa/US11_NEW_PUB.pap:*
8: /SIDS5/ptodata/1/pubppa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	276.5	9.0	802	US-10-510-386-2	Sequence 2, Appl1
2	274.5	9.0	874	US-10-510-386-28	Sequence 28, Appl1
3	274.5	9.0	1047	US-10-510-386-200	Sequence 200, Appl1
4	249.5	8.2	379	US-11-156-062-23	Sequence 23, Appl1
5	242	7.9	272	US-11-020-602-236	Sequence 236, Appl1
6	239	7.8	269	US-11-020-602-6	Sequence 6, Appl1
7	237	7.7	382	US-11-020-602-2	Sequence 2, Appl1
8	227.5	7.4	274	US-11-156-062-14	Sequence 14, Appl1
9	227.5	7.4	275	US-11-065-943-54	Sequence 54, Appl1
10	227.5	7.4	275	US-11-020-602-3	Sequence 3, Appl1
11	226.5	7.4	274	US-11-156-062-12	Sequence 12, Appl1
12	225.5	7.4	274	US-11-020-602-5	Sequence 5, Appl1
13	223.5	7.3	274	US-11-156-062-4	Sequence 4, Appl1
14	221.5	7.2	274	US-11-156-062-10	Sequence 10, Appl1
15	220.5	7.2	274	US-11-156-062-6	Sequence 6, Appl1
16	220.5	7.2	275	US-11-156-062-8	Sequence 8, Appl1
17	217.5	7.1	274	US-11-020-602-4	Sequence 4, Appl1
18	217.5	7.1	274	US-11-156-062-16	Sequence 16, Appl1
19	217.5	7.1	274	US-11-156-062-18	Sequence 18, Appl1
20	216.5	7.1	274	US-11-156-062-2	Sequence 2, Appl1
21	210.5	6.9	1052	US-11-020-602-208	Sequence 208, Appl1
22	204.5	6.7	1432	US-10-510-386-218	Sequence 218, Appl1
23	198.5	6.5	280	US-11-020-602-209	Sequence 209, Appl1
24	197.5	6.5	1647	US-11-052-554A-260	Sequence 260, Appl1
25	179	5.9	740	US-11-096-568A-24714	Sequence 24714, A

26	179	5.9	777	US-11-096-568A-24713	Sequence 24713, A
27	179	5.9	790	US-11-096-568A-24712	Sequence 24712, A
28	166.5	5.4	591	US-10-510-386-22	Sequence 22, Appl1
29	160	5.2	722	US-11-096-568A-31863	Sequence 31863, A
30	160	5.2	757	US-11-096-568A-31862	Sequence 31862, A
31	160	5.2	798	US-11-096-568A-31861	Sequence 31861, A
32	152	5.0	733	US-11-096-568A-24028	Sequence 24028, A
33	152	5.0	759	US-11-096-568A-24027	Sequence 24027, A
34	152	5.0	764	US-11-096-568A-24026	Sequence 24026, A
35	151	4.9	659	US-11-096-568A-17896	Sequence 17896, A
36	151	4.9	791	US-11-096-568A-17895	Sequence 17895, A
37	151	4.9	791	US-11-096-568A-31251	Sequence 31251, A
38	150	4.9	791	US-11-096-568A-31251	Sequence 31251, A
39	148.5	4.9	794	US-11-218-986-2	Sequence 2, Appl1
40	148.5	4.9	820	US-10-821-234-1176	Sequence 1176, Appl1
41	143	4.7	672	US-11-096-568A-23983	Sequence 23983, A
42	143	4.7	680	US-11-096-568A-23982	Sequence 23982, A
43	134	4.4	470	US-10-873-528-3	Sequence 3, Appl1
44	132.5	4.3	617	US-11-096-568A-31252	Sequence 31252, A
45	130	4.2	418	US-11-096-568A-23984	Sequence 23984, A

ALIGNMENTS

```
RESULT 1
US-10-510-386-2
; Sequence 2, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjørke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-2

Query Match          9.0% Score 276.5; DB 6; Length 802;
Best Local Similarity 21.8% Pred. No. 8.5e-14;
Matches 160; Conservative 82; Mismatches 226; Indels 265; Gaps 31;

QY 126 DVKSKXXIIEHVESVEPYLPYXXIDPOLFTGASXLYVAXALDT---KQXNKEVOLRGIE 182
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 49 DVTSPKQTTTVVELKESLAKADGKQTKASLTKASKYLTKAKVAREVD-RVPS 107
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 183 XIAQXXNSNDVXYITAKPERYVM-----NDVARGIVRAD 216
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 108 GFSMKLPASEIRPKLALAVEKAVVFNATYKPDVSXGKQVTLAADAIVPQMKASAPFICAD 167
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 217 VAQSSYGLYGQGVAVADTGLDTCGRND--SSMHEAFRKK--ITLALAGRTNNADTNG- 272
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 168 QAWKS-GTGTGKIKVAVLDTVDYTHPLKKNFGPKYDVPDNDYDPOETPTTGPGRGA 226
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 273 --HGTHTVAGSVLGNKXTKNGAPQNALVFQSIIMDSXGGLPSNLOTLFQOAXSAGARI 330
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 227 TDHGHVHVGITLAANGQI-KGVAPRATLILAVVLGRG--SGTTENVIGIEKAVADGAKV 283
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 331 HTNKGAAVNGA-YTTDSRNVDYVRKNDITLLPAAGEXKNGGTISAPGAKAIVTGA 389
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 284 MNLSTGKSLNSPDVAISI--ALDWAMEGVAVLVISNGSGENMTVCSPTGSRVALSVGA 341
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 390 TE-----NLRPSRGS----- 399
    |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```
Db 342 SQLPYNEXSVTLPSYSAKVMGVOEBKOLEALNGQVELVEAGLQADDPFGSKGVKGYA 401
Qy 400 -----YADNINH----- 406
Db 402 VIQGVIPFVDKAEKAKAGALGAVIYNNATGEIBANVMGMAVPTVKLSKEBGEKLVQOI 461
Qy 407 -----VAQSSRGPTKO-GRIKPDVMAFGTXIISARSLSLPDSSF 445
Db 462 KEKGKSVSFFLDKCKLGETTISFSSRGFPVMDTMMIKPDVSAFGVNIYSTIPT----- 514
Qy 446 WANHDSK---YAVWGTSMATPIVAGNVAQLREHFVKNRGITP--KPSLLKALILAGNA 499
Db 515 ---HDPKAPYGVGSGQGTSMASPHVAGTAAILKQ-----AKPDWTEQIIGVLMNTAB 564
Qy 500 ---DXGLGYPNGNGKRGVTLDSKLVN-----AYNNESS 530
Db 565 KLTBNGPRLPHNTQAGSIRIMEALKASIVTPGSHSVGTPLKDKGQKQKKAFTIEN- 623
Qy 531 XLSTSQKA---TYXFTATAGKPLKISLWSD--APASTT---ASVTLVNDLIVITAPN 581
Db 624 -LSHRKAIYQLEYSFKGTG-----ITVSGTERVVPVPAQGTGAALKVTVNSAKTAYE 677
Qy 582 GTXYVGNDFKXXPKXXNMDGRNVE--NVFINKP-----QSGTYTIEVOAY 624
Db 678 GTVYIRB-----DQRKVAEIRPLLIVKEPDYPRVTSVTVBPAKQAGATTB--AY 725
Qy 625 NVFVPGFQKFSIAI 637
Db 726 -LPGAEELAFV 737

RESULT 2
US-10-510-386-28
; Sequence 28, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 28
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-28

Query Match 9.0%; Score 274.5; DB 6; Length 874;
Best Local Similarity 23.7%; Pred. No. 1.4e-13;
Matches 141; Conservative 85; Mismatches 253; Indels 117; Gaps 25;

Qy 13 AAILSTVALXNPSAGAKRFPDLDFKGIQTDDXXGFSKQQTGAAP-----LV 62
Db 213 AAPISAKMLFNKSRDSDVYKDL---VQLKPLFADVAKQGVASITNDQKASISLY 268
Qy 63 ESENVKXKXGKXKLETPANNKLIHQPNPILBETKQILEXG-AKILDIYIPDIYAIY 121
Db 269 ETAPASVEPLKQQLDQVAKD--IGIBQLTG---SKVASVLEKAGMATASSAPENRIV 323
Qy 122 EYE-----GDVXSXXXXIEHVESVEP-----YLPYXID-POLFTGASXLVKAHALD 168
Db 324 KXGKXKRGFSKSKAQS--SGVQALEPLGKSKTAFKMTVYEMKESRSGFRAAKQYQAA 382
Qy 169 TKQXKKEVQLRGIEIXIAQ-XXXXSDVXYITAKPEYKVNADVARGIVKADVAQSSYGLYQ 227
Db 383 ASKIAKMEVEFEVQVQYEAALSRDTQY---PYQWSLKNNGNKRAANADIQFEOQLKLMK 439
```

```
Qy 228 GQ-----IVAVADTGLDGRNDSMHEAFRGKI---TALVALGRTNANDNGHGTAVG 279
Db 440 GKXKDTVIADVDTGVDBHTLADLS-----GSVKDREGYNTVGRPADMDNDNGHGTAVSG 493
Qy 280 SVLG---NGKTNKGAAPQANLVFOSIMDSXGGLGSLPSNIQTLFPOAKXSAGARIHTNSWG 336
Db 494 IIAAQNDFHFMAGINAVAKILPVKVLDDSG--SGDPEQIANGIIVAADHAKYINLSLG 551
Qy 337 AAVNGAYTDSRNVDDYRKQDMTILFPAQHEKPNNGGTISAPGTAKAIIYGAETENTRPS 396
Db 552 ---GPYSRWMEYALKTAASKNVTIYVAATGNDGVS--EISYPASSKTYLSGATNNL--- 602
Qy 397 FGSAVDNINHQAQPSRGPTDGRIKPDVMAFGTXIISARSLSLPDSSFMANHDSKXAYM 456
Db 603 -----DLVSDYSNKGKGI-----DMVAPGTDI---PSLVPDGN-----VTYM 636
Qy 457 GGTSMATPIVAGNAQLREHFVKNRGITPKEPSLLKALILAGNADXL---GYPNGN--- 509
Db 637 SGTSMAPVHVAAGGL-----LSQNPISLKRQJASILTEHTADVAPEBQDNRPDPDLD 691
Qy 510 -----QKGRVTLDSKLVNAYNNESSXLSSTQKATYXFTATGAKPLKI 552
Db 692 IEPAAQIPGYDFVSGWGLNVFHAASVFEIEMKVHPVLNRTAVTGTAKSGVTVKI 747

RESULT 3
US-10-510-386-200
; Sequence 200, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 200
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-200

Query Match 9.0%; Score 274.5; DB 6; Length 1047;
Best Local Similarity 23.7%; Pred. No. 1.7e-13;
Matches 141; Conservative 85; Mismatches 253; Indels 117; Gaps 25;

Qy 13 AAILSTVALXNPSAGAKRFPDLDFKGIQTDDXXGFSKQQTGAAP-----LV 62
Db 213 AAPISAKMLFNKSRDSDVYKDL---VQLKPLFADVAKQGVASITNDQKASISLY 268
Qy 63 ESENVKXKXGKXKLETPANNKLIHQPNPILBETKQILEXG-AKILDIYIPDIYAIY 121
Db 269 ETAPASVEPLKQQLDQVAKD--IGIBQLTG---SKVASVLEKAGMATASSAPENRIV 323
Qy 122 EYE-----GDVXSXXXXIEHVESVEP-----YLPYXID-POLFTGASXLVKAHALD 168
Db 324 KXGKXKRGFSKSKAQS--SGVQALEPLGKSKTAFKMTVYEMKESRSGFRAAKQYQAA 382
Qy 169 TKQXKKEVQLRGIEIXIAQ-XXXXSDVXYITAKPEYKVNADVARGIVKADVAQSSYGLYQ 227
Db 383 ASKIAKMEVEFEVQVQYEAALSRDTQY---PYQWSLKNNGNKRAANADIQFEOQLKLMK 439
Qy 228 GQ-----IVAVADTGLDGRNDSMHEAFRGKI---TALVALGRTNANDNGHGTAVG 279
Db 440 GKXKDTVIADVDTGVDBHTLADLS-----GSVKDREGYNTVGRPADMDNDNGHGTAVSG 493
Qy 280 SVLG---NGKTNKGAAPQANLVFOSIMDSXGGLGSLPSNIQTLFPOAKXSAGARIHTNSWG 336
```

Db 494 IIAAANDHFMAGINAVAKILEPVKYLDSG--SGDTEQIANGIIYAADHGAKVINLSIG 551
Qy 337 AAVNGAYTTDSRNVDDYVAKNDMTILFAAGNEXPNCGTISAPCTAGATVAGTENTLRPS 396
Db 552 ---GYSVMEYALRYAASKNVTIYAAGNDGVS--EISYPASSKYTISVGAITNVL--- 602
Qy 397 FCSYADINNHVAQFSSRGFTKQGRIKPDYMAPGTXTLSARSSSLAPDSSFANHDSKAYM 456
Db 603 ---DVSVDYNYGKGL-----DVAHPTD-----PSLVPGN-----VTTM 636
Qy 457 GGTSMATPIVAGNVAQREHFVNKRGITPKPSILKAALJAGADKGL---GYPNGN--- 509
Db 637 SGTSMAPVHVAAGL-----LSQNPSPKPKQIASLITETPADVAFBEGDNNPNPDYDD 691
Qy 510 -----OCGRVTLDSKSLNVAAYNNESSKLSSTQKATYXTATNAPKPLKI 552
Db 692 IEPAAQIPGYDFVSGWGLNVPFAASVFELNMKVHPVLRHRAVGTAKSGVTVKI 747

RESULT 4

US-11-156-062-23
; Sequence 23, Application US/11156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stenr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / HS698
; CURRENT APPLICATION NUMBER: US/11/156,062
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 379
; TYPE: PRY
; ORGANISM: Bacillus licheniformis
US-11-156-062-23

Query Match 8.2%; Score 249.5; DB 7; Length 379;
Best Local Similarity 23.7%; Pred. No. 4.1e-12;
Matches 117; Conservative 58; Mismatches 164; Indels 155; Gaps 18;

Qy 1 MRKKVFLSVLSAAAILSTVALXNPSAG--ARXFDLPKGIQTDDXKGFSGKQQTGA 58
Db 2 MRKKSFWLGLTAFMLVFTMAPSDSASAQPAKNVEKDY-----YGFSGVATAS- 52
Qy 59 AFLVESENVKLGKGLKLETPANNKLIHQFNGPILEETKQXLEXTAKLIDYIPDA 118
Db 53 ---VKDIIKESGGKVDK-----QFR--IYAAAKAKDKALKKVKNDPDVA 94
Qy 119 YIVVEGVDVXXXXIENHVSVEPYLPXYXIDPOLFTKGSXLYVAKXALDTQXNKEVQL 178
Db 95 YVEEY-----DHVAHALAQTVPGIP----- 114
Qy 179 RGIEKIAQXXXNDVXYITAKEPYKMDVAGIYKADVAQSSYGLYGGQGIYVAVDTGL 238
Db 115 -----LTKADRYQAQ--GFKGANVAVVAVLDTGI 140
Qy 239 DTGRNDSVHBAFRGKITVALGRTNNANDTNGHTVAGSV--LGNXTNKGAPOAN 296
Db 141 QASHPDANVVG-----ASPTVAGKAYN--TDGNGHTVAGVVAALDNTTGVLGVAPSVS 193
Qy 297 LVFQSIQDSXGGLGILPSNLQTLFQOAKSAGARIHTNSGAAVNGAYTTDSRNVDDYAK 356
Db 194 LYAVKVLNSSG--SGSVSGIVSGIEMATNTGMDVIMNSIGGA---SGSTAMKQAVDNVA 248

Qy 357 NDMTILFAAGNEXPNCG--GTISAPGTAKNATVAGTENTLRPSFGSYADINNHVAQFSSRG 414
Db 249 KGVVVAAGNCGSGGNTTITIGYPAKTDVIAVGA-----VDSNSRASFSSVG 297
Qy 415 PYKDGRIKPDVAPGTXTLSARSSSLAPDSSFANHDSKAYMGSTMATPIVAGNVA--- 471
Db 298 -----ALEVNAFPGAGVSTYPT-----NTYATLNGTSMASPHVAGAAALIL 339
Qy 472 ---QREHFVZOR 481
Db 340 SKHPNLSAQVNR 353

RESULT 5

US-11-020-602-236
; Sequence 236, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Batelli, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GCS27C2
; CURRENT APPLICATION NUMBER: US/11/020,602
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 236
; LENGTH: 272
; TYPE: PRY
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hybrid of
; OTHER INFORMATION: Bacillus lencus and Bacillus amyloliquefaciens
US-11-020-602-236

Query Match 7.9%; Score 242; DB 7; Length 272;
Best Local Similarity 31.9%; Pred. No. 1e-11;
Matches 84; Conservative 30; Mismatches 101; Indels 48; Gaps 10;

Qy 213 VQADVAQSSYGLYGGQGIYVAVDTGLDTGRNDSVHBAFRGKITVALGRTNNANDTNG 272
Db 11 VQAPAAHNR--GLTSGGVVAVLDTGIST-----HPDLNIRGASFPVGBP--STQDNG 61
Qy 273 HGTHTVAGSV--LGNXTNKGAPOANLVFQSIQDSXGGLGILPSNLQTLFQOAKSAGARI 330
Db 62 HGTHTVAGTIALMNSIGVIGVAPSAELVAVKVLGASG--SGSVSSIAQGLEWAGNNGMAY 119
Qy 331 HTNSGAAVNGAYTTDSRNVDDYVAKNDMTILFAAGNEXPNCG--GTISAPGTAKNATVAG 388
Db 120 IMNSLGGSGAAL---KAADKAVASGVVVAAGNCGTSSSTVIGFPGRTVPIAVG 175
Qy 389 ATENLRPSFGSYADINNHVAQFSSRGFTKQGRIKPDVAPGTXTLSARSSSLAPDSSFAN 448
Db 176 A-----VDSNQRASFSSVGF-----ELDVAVAPGVISIGSTLPG----- 208
Qy 449 HDSKTAVMGSTMATPIVAGNVA 471
Db 209 --NKYGAVNGTSMASPHVAGAA 229

RESULT 6

US-11-020-602-6
; Sequence 6, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Batelli, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND


```

RESULT 8
US-11-156-062-14
; Sequence 14, Application US/11156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Mieland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueger, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; PRIOR FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-14

```

```

Query Match      7.4%; Score 227.5; DB 7; Length 274;
Best Local Similarity 30.2%; Pred. No. 1.4e-10;
Matches 85; Conservative 35; Mismatches 108; Indels 53; Gaps 11;

```

```

QY 212 YKADVAQSSYGLYGGQCIYAAVADTGLDGRNDSMHEAFRGKITLALYALGRTNNANDTN 271
DB 10 LKAP-KVQAQ-GFKGANVVAALDTGICQASHPLNVGG-----ASFVAGEAYNA-DDN 61
QY 272 HGTHVAGSV--LGNGXTNKGMAPQANLVFQSIIMDSXGGLGSPNLQTLFSGQXSAGAR 329
DB 62 HGTHVAGVAAALDNTTGTGLVAFPSVSLVAVKVLSNG--SSYSIGIVSGIEWATTNGMD 119
QY 330 HTNSMGAAVNGAYTTDSRNVDDYRKNDMTILFAAGNEXPN--GTISAPGTAKNAITV 387
DB 120 VINMSLGA---SGSTAMKQAVDNAAVAGVVAVAAAGSSGNTTITGPAKYDSVIAV 176
QY 388 GATENLRPSFGSYADNINNVAFQSSRGPTKDGRIKPDVMAFGTIIISARSSLAPDSSFWA 447
DB 177 GA-----VDSNSRASFSSVG-----AELFVAPGAGVSTYPT----- 210
QY 448 HNSKXYVMGSTMATPIVAGNVA-----QLRHFVKNR 481
DB 211 ---NTYATWNGTSMASPHVAGAAALILSKHPNLASQVNR 248

```

```

RESULT 9
US-11-065-943-54
; Sequence 54, Application US/11065943
; Publication No. US20050250131A1
; GENERAL INFORMATION:
; APPLICANT: JESTIN, JEAN-LUC
; APPLICANT: VICHIER-GUERRE, SOPHIE
; APPLICANT: FERRIS, STREPHANE
; TITLE OF INVENTION: METHODS FOR OBTAINING THERMOSTABLE ENZYMES, DNA POLYMERASE I
; TITLE OF INVENTION: VARIANTS FROM THERMUS AQUATICUS HAVING NEW CATALYTIC ACTIVITIES,
; FILE REFERENCE: 266426USOXCI
; CURRENT APPLICATION NUMBER: US/11/065,943
; PRIOR FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: US 10/787,219
; PRIOR FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 54
; LENGTH: 275

```

```

; TYPE: PRT
; ORGANISM: Bacillus amyloliquefaciens
US-11-065-943-54

```

```

Query Match      7.4%; Score 227.5; DB 7; Length 275;
Best Local Similarity 28.7%; Pred. No. 1.4e-10;
Matches 81; Conservative 26; Mismatches 92; Indels 83; Gaps 9;

```

```

QY 213 YKADVAQSSYGLYGGQCIYAAVADTGLDGRNDSMHEAFRGKITLALYALGRTNNANDTN 272
DB 11 IKAP-ALHSQGYTGSNVKAVVIDSGID-----SSHPLKAVAGASMPVSEETPFQDNNS 63
QY 273 HGTHVAGSV--LGNGXTNKGMAPQANLVFQSIIMDSXGGLGSPNLQTLFSGQXSAGAR 330
DB 64 HGTHVAGVAAALNNSIGVLGVAAPSASLVAVKVLGADG----- 100
QY 331 HTNSMGAAVNGAYTTDSRNVDDYRKNDMTILFAAGNEXPN 371
DB 101 -SGQYSWIINGIEMALANNMDVINMSLGPSGSAALKAADVAGVAVVAAAGNEGTS 159
QY 372 G--GTISAPGTAKNAITVGATENLRPSFGSYADNINNVAFQSSRGPTKDGRIKPDVMAFG 429
DB 160 GSSSTVGYPGKYPVIAVGA-----VDSNQRPASFSSVGP-----ELDVMAFG 202
QY 430 TXILSARSSLAPDSSFWAHNSKXYVMGSTMATPIVAGNVA 471
DB 203 VSIOSTLPG-----NKYAGYNGTSMASPHVAGAAA 232

```

```

RESULT 10
US-11-020-602-3
; Sequence 3, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Bateall, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GCS27C2
; CURRENT APPLICATION NUMBER: US/11/020,602
; PRIOR FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Bacillus amyloliquefaciens
US-11-020-602-3

```

```

Query Match      7.4%; Score 227.5; DB 7; Length 275;
Best Local Similarity 28.7%; Pred. No. 1.4e-10;
Matches 81; Conservative 26; Mismatches 92; Indels 83; Gaps 9;

```

```

QY 213 YKADVAQSSYGLYGGQCIYAAVADTGLDGRNDSMHEAFRGKITLALYALGRTNNANDTN 272
DB 11 IKAP-ALHSQGYTGSNVKAVVIDSGID-----SSHPLKAVAGASMPVSEETPFQDNNS 63
QY 273 HGTHVAGSV--LGNGXTNKGMAPQANLVFQSIIMDSXGGLGSPNLQTLFSGQXSAGAR 330
DB 64 HGTHVAGVAAALNNSIGVLGVAAPSASLVAVKVLGADG----- 100
QY 331 HTNSMGAAVNGAYTTDSRNVDDYRKNDMTILFAAGNEXPN 371
DB 101 -SGQYSWIINGIEMALANNMDVINMSLGPSGSAALKAADVAGVAVVAAAGNEGTS 159
QY 372 G--GTISAPGTAKNAITVGATENLRPSFGSYADNINNVAFQSSRGPTKDGRIKPDVMAFG 429
DB 160 GSSSTVGYPGKYPVIAVGA-----VDSNQRPASFSSVGP-----ELDVMAFG 202

```

Qy	430	TXILSRRSLAPDSSFWMANHDSKYAMGCTSMATPIVAGNVA	471
		: :	
Db	203	VSIQSTLP-----NRYGAYNGTSMASPHVAGAA	232

RESULT 11

```

US-11-156-062-12
: Sequence 12, Application US/11156062
: Publication No. US20050281773A1
: GENERAL INFORMATION:
: APPLICANT: Wieland, Susanne
: APPLICANT: Polanyi-Bald, Laura
: APPLICANT: Prueser, Inken
: APPLICANT: Stehr, Regina
: APPLICANT: Maurer, Karl-Heinz
: TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PEPHYDROLASE ACTIVITY
: FILE REFERENCE: HEKK-0134 / H5698
: CURRENT APPLICATION NUMBER: US/11/156, 062
: CURRENT FILING DATE: 2005-06-17
: PRIOR APPLICATION NUMBER: PCT/EP2003/014127
: PRIOR FILING DATE: 2003-12-20
: PRIOR APPLICATION NUMBER: DE 102 60 903.9
: PRIOR FILING DATE: 2002-12-20
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 12
: LENGTH: 274
: TYPE: PRN
: ORGANISM: Artificial
: FEATURE:
: OTHER INFORMATION: Synthetic Construct
US-11-156-062-12

```

Query Match	7.48; Score 226.5; DB 7; Length 274;

Qy	212	1VKADVAOSYGLVGGQGVVAADVTGLDGTGRNDSMHEAFRGKTALYALGTTNANDTN	271
		:: ::	
Db	10	LKKADPKVQAQ-GFKKANVKVALLDGLGASHDDLNVVVG-----ASFYAGRAYNA-DGN	61
Qy	272	GHGTHVAGSV--LGNCGTNNKGMAPOANLVFQSIMDSGGGLGGLPSNLQTLVFSQAXSAGAR	329
Db	62	GHGTHVAGTVAALDNTTGLVGVAPSVSVAYAVVLDSSG--SSYSYGIVSGIMATTNGMD	119
Qy	330	IHTNSKGAAVNGAVYTTDSRNVDDYVRKQDMTLLFPAENGKPNQ--GTISAPGTAKNAITV	387
		:: ::	
Db	120	VINNSLGA---SGSTAMKQAVDNAYARGVVVAAAGNSGSSGNTNTGYPAKYDSVIAV	176
Qy	388	GATENLPLSPFSYADNINHVAQFSSRGPTKGRIRIDYDVAAPPTXLLASRSLAPDSSFWA	447
		:: ::	
Db	177	GA-----VDSNSNRASFSSVQ-----AELEVAAPAGVYSTPT-----	210
Qy	448	NHDSKYAVMGSTMAPTPIVAGNVA-----QLREHFYKNR	481
Db	211	---NTVATNGTSMASPHVAGAAALILSKHPILSASQVKNR	248

RESULT 12
US-11-020-602-5

```

; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.1

```

Query Match 7.4%; Score 225.5; DB 7; Length 274;
Best Local Similarity 29.9%; Pred. No. 2e-10;
Matches 84; Conservative 36; Mismatches 108; Indels 53; Gaps 11;

[illegible]

```

RESULT 13
US-11-156-062-4
; Sequence 4, Application US/11156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; US-11-156-062-4

```

Query Match	7.3%;	Score 222.5;	DB 7;	Length 274;
Best Local Similarity	29.9%;	Pred. No. 3.5e-10;		
Matches	84;	Conservative 35;	Mismatches 109;	Indels 53;
				Gaps 11;

QY	212	IIVADVAAGSYGLGCGQIYAVALDTGLDTPGRDNSSMHEARKKITALYALGRITNADPTN	271
	:	: :: :	:
D8	10	LIIADKKVQAO-GRKANAVKVALDGTQASHPDILNVGG-----LSFVAGEAYN-IDGN	61
	:	: :: :	:
QY	272	GHCITHVAGSV--LGNGXTNKGMAPLANLVEQSIMDSXGIGLGPLSNLTQTFPQNXSGAR	329
	:	: :: :	:
b6	62	GCHTHVAGTAAALDNTGTGVAPPSVLVYVKRLNSGG--SGXSXGIGVGIEAMATTKGMD	119
	:	: :: :	:

```

QY 330 IHTNSGAAVNGAYTTDSRNVDYVRKNDMTILFAAGNEXPNG--GTISAPGTAKNAITV 387
DB 120 VINMSLGGA---SGSTAKKQAVDNAYARGVVVAAAGNCGSSGNTTIGYPAKYSVIAV 176
QY 388 GATENLRPSFGSYADNINHVAQFSSRGPTKGRIRKPDVAPGTXTILSARSLAPDSSFMA 447
DB 177 GA-----VDSNSNRASFSVSG-----AELBWAAPGAGVYSTYPT----- 210
QY 448 NHDSKYAMGCTSMATPIYAGNVA-----QLRHFYKXR 481
DB 211 ---NTYATWDTGTSMAFPHVAGAAALLILSKHPILSASQVNR 248

```

RESULT 14

```

US-11-156-062-10
; Sequence 10, Application US/11156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stenif, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-10

```

```

Query Match 7.2%; Score 221.5; DB 7; Length 274;
Best Local Similarity 29.9%; Pred. No. 4,2e-10;
Matches 84; Conservative 35; Mismatches 109; Indels 53; Gaps 11;

```

```

QY 212 IYKADVAQSSYGLYQGGIYAVADTGLDGRNDSMHEAFRGKITALYALGRTNNANDTN 271
DB 10 LIKADKVOAQ-GFKGANVVAVLDTGIGQSHPDILNVVG-----ASFVAGRAYNA-DGN 61
QY 272 GHGTVAGSV--LNGKXTKKGMAPOANLVFOSIMDSXGGLGSLPSNLQTLFPOAKSAGAR 329
DB 62 GHGTVAGTVALLDITGVLGVAFVSYSYAVKVLNSSG--SGSYSIGIVSGIEMATTNGMD 119
QY 330 IHTNSGAAVNGAYTTDSRNVDYVRKNDMTILFAAGNEXPNG--GTISAPGTAKNAITV 387
DB 120 VINMSLGGA---SGSTAKKQAVDNAYARGVVVAAAGNCGSSGNTTIGYPAKYSVIAV 176
QY 388 GATENLRPSFGSYADNINHVAQFSSRGPTKGRIRKPDVAPGTXTILSARSLAPDSSFMA 447
DB 177 GA-----VDSNSNRASFSVSG-----AELBWAAPGAGVYSTYPT----- 210
QY 448 NHDSKYAMGCTSMATPIYAGNVA-----QLRHFYKXR 481
DB 211 ---NTYATWDTGTSMAFPHVAGAAALLILSKHPILSASQVNR 248

```

```

RESULT 15
US-11-156-062-6
; Sequence 6, Application US/11156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne

```

```

; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stenif, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-6

```

```

Query Match 7.2%; Score 220.5; DB 7; Length 274;
Best Local Similarity 29.9%; Pred. No. 5e-10;
Matches 84; Conservative 35; Mismatches 109; Indels 53; Gaps 11;

```

```

QY 212 IYKADVAQSSYGLYQGGIYAVADTGLDGRNDSMHEAFRGKITALYALGRTNNANDTN 271
DB 10 LIKADKVOAQ-GFKGANVVAVLDTGIGQSHPDILNVVG-----ASFVAGRAYNA-DGN 61
QY 272 GHGTVAGSV--LNGKXTKKGMAPOANLVFOSIMDSXGGLGSLPSNLQTLFPOAKSAGAR 329
DB 62 GHGTVAGTVALLDITGVLGVAFVSYSYAVKVLNSSG--SGSYSIGIVSGIEMATTNGMD 119
QY 330 IHTNSGAAVNGAYTTDSRNVDYVRKNDMTILFAAGNEXPNG--GTISAPGTAKNAITV 387
DB 120 VINMSLGGA---SGSTAKKQAVDNAYARGVVVAAAGNCGSSGNTTIGYPAKYSVIAV 176
QY 388 GATENLRPSFGSYADNINHVAQFSSRGPTKGRIRKPDVAPGTXTILSARSLAPDSSFMA 447
DB 177 GA-----VDSNSNRASFSVSG-----AELBWAAPGAGVYSTYPT----- 210
QY 448 NHDSKYAMGCTSMATPIYAGNVA-----QLRHFYKXR 481
DB 211 ---NTYATWDTGTSMAFPHVAGAAALLILSKHPILSASQVNR 248

```

```

Search completed: April 7, 2006, 18:04:19
Job time : 17.4871 secs

```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 17:58:27 ; Search time 118.593 Seconds
(without alignments)
2254.868 Million cell updates/sec

Title: US-10-784-870-2

Perfect score: 3059
Sequence: 1 MRXKKVFLSVLSAAALIST.....EVQATNVGVGQXFLSATVN 640

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3029	99.0	640	US-09-920-954-6	Sequence 6, App1
2	3029	99.0	640	US-10-456-479-6	Sequence 4, App1
3	3029	99.0	640	US-10-784-870-6	Sequence 4, App1
4	3029	99.0	640	US-10-820-712A-3	Sequence 3, App1
5	3029	99.0	640	US-10-820-714A-3	Sequence 3, App1
6	3028	99.0	640	US-09-920-954-8	Sequence 8, App1
7	3028	99.0	640	US-10-784-870-8	Sequence 8, App1
8	3020	98.7	639	US-09-920-954-4	Sequence 4, App1
9	3020	98.7	639	US-10-784-870-4	Sequence 4, App1
10	3016	98.6	640	US-09-920-954-2	Sequence 2, App1
11	3016	98.6	640	US-10-784-870-2	Sequence 2, App1
12	3007	98.3	639	US-09-920-954-1	Sequence 1, App1
13	3007	98.3	639	US-10-784-870-1	Sequence 1, App1
14	2155	70.4	434	US-09-985-689A-1	Sequence 1, App1
15	2155	70.4	434	US-09-985-689A-2	Sequence 2, App1
16	2155	70.4	434	US-10-385-662-2	Sequence 2, App1
17	2155	70.4	434	US-10-456-479-2	Sequence 2, App1
18	2155	70.4	434	US-10-456-479-10	Sequence 10, App1
19	2155	70.4	434	US-10-456-479-11	Sequence 11, App1
20	2155	70.4	434	US-10-837-566-1	Sequence 1, App1
21	2155	70.4	434	US-10-837-566-2	Sequence 2, App1
22	2155	70.4	434	US-10-820-712A-1	Sequence 1, App1
23	2155	70.4	434	US-10-820-712A-12	Sequence 12, App1
24	2155	70.4	434	US-10-820-714A-14	Sequence 14, App1
25	2155	70.4	434	US-10-820-714A-1	Sequence 1, App1
26	2155	70.4	434	US-10-820-714A-13	Sequence 13, App1
27	2155	70.4	434	US-10-820-714A-15	Sequence 15, App1

28	2082	68.1	434	US-09-985-689A-6	Sequence 6, App1
29	2082	68.1	434	US-10-456-479-15	Sequence 15, App1
30	2082	68.1	434	US-10-837-566-6	Sequence 6, App1
31	2082	68.1	434	US-10-820-712A-22	Sequence 22, App1
32	2082	68.1	434	US-10-820-714A-23	Sequence 23, App1
33	2060.5	67.4	433	US-09-985-689A-7	Sequence 7, App1
34	2060.5	67.4	433	US-10-456-479-16	Sequence 16, App1
35	2060.5	67.4	433	US-10-837-566-7	Sequence 7, App1
36	2060.5	67.4	433	US-10-820-712A-23	Sequence 23, App1
37	2060.5	67.4	433	US-10-820-714A-24	Sequence 24, App1
38	1952.5	63.8	433	US-09-985-689A-5	Sequence 5, App1
39	1952.5	63.8	433	US-10-456-479-14	Sequence 14, App1
40	1952.5	63.8	433	US-10-837-566-5	Sequence 5, App1
41	1952.5	63.8	433	US-10-820-712A-20	Sequence 20, App1
42	1952.5	63.8	433	US-10-820-714A-21	Sequence 21, App1
43	1948.5	63.7	433	US-09-985-689A-3	Sequence 3, App1
44	1948.5	63.7	433	US-10-456-479-12	Sequence 12, App1
45	1948.5	63.7	433	US-10-837-566-3	Sequence 3, App1

ALIGNMENTS

RESULT 1					
US-09-920-954-6					
Sequence 6, Application US/09920954					
GENERAL INFORMATION:					
APPLICANT: TAKAIIWA, MIKIO					
APPLICANT: OKUDA, MITSUYOSHI					
APPLICANT: SAKETI, KATSUHIISA					
APPLICANT: KUBOTA, HIROMI					
APPLICANT: HITOMI, JUN					
APPLICANT: KAGEYAMA, YASUSHI					
APPLICANT: SHIKATA, SHITSUM					
APPLICANT: NOMURA, MASAFUMI					
TITLE OF INVENTION: ALKALINE PROTEASE					
FILE REFERENCE: 0327-0832-0PCT					
CURRENT APPLICATION NUMBER: US/09/920,954					
PRIOR FILING DATE: 2001-08-03					
PRIOR APPLICATION NUMBER: 09/509,814					
PRIOR FILING DATE: 2000-04-06					
PRIOR APPLICATION NUMBER: PCT/JP98/04528					
PRIOR FILING DATE: 1998-10-07					
PRIOR APPLICATION NUMBER: JP 9-274570					
PRIOR FILING DATE: 1997-06-08					
NUMBER OF SEQ ID NOS: 24					
SOFTWARE: PatentIn version 3.0					
SEQ ID NO 6					
LENGTH: 640					
TYPE: PRT					
ORGANISM: Bacillus sp.					
US-09-920-954-6					
Query Match					
Best Local Similarity 93.3%; Pred. No. 8.7e-26;					
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;					
QY	1	MRXKKVFLSVLSAAALISTVAAKNPSAGXARPDLDKGIQTTDDXXGFSKXOTGAA	60		
DB	1	MRXKKVFLSVLSAAALISTVAAKNPSAGXARPDLDKGIQTTDDXXGFSKXOTGAA	60		
QY	61	FLVSENVKLLKGLKKLETPANNKKHIXQFNGPILEBTRQXLEXTGAKTLDYIPDAY	120		
DB	61	FLVSENVKLLKGLKKLETPANNKKHIXQFNGPILEBTRQXLEXTGAKTLDYIPDAY	120		
QY	121	IVRYGVDVKSATSTEHVSVEPYLPYRIDPQLETKASBLVAVVADDTQKNKEVQLR	180		
DB	121	IVRYGVDVKSATSTEHVSVEPYLPYRIDPQLETKASBLVAVVADDTQKNKEVQLR	180		
QY	181	GIEIXAIXXKXNDVYITAKPEYKAMDVAGIYKADVAOSYGLYGGQIVAVADTGLD	240		
DB	181	GIEIXAIXXKXNDVYITAKPEYKAMDVAGIYKADVAOSYGLYGGQIVAVADTGLD	240		

QY 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGXTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGXTNKGMAPOANLVF 300
QY 301 QSIMDSXGGLGGLPSNLQTLFQOAXSAGARLHTNSWGAUVNGAYTTDSRNVDYVRKNDM 360
DB 301 QSIMDSXGGLGGLPSNLQTLFQOAXSAGARLHTNSWGAUVNGAYTTDSRNVDYVRKNDM 360
QY 361 TILPAAGNEXPNNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRPTKDG 420
DB 361 TILPAAGNEXPNNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRPTKDG 420
QY 421 RIKPDVMAPGTIIISARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
DB 421 RIKPDVMAPGTIIISARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
QY 481 NRGITPKPSLKAALIAAGADIGLGYPNGNGMGRVTLDKSLNVAAYNESSLSLSQKAT 540
DB 481 NRGITPKPSLKAALIAAGADIGLGYPNGNGMGRVTLDKSLNVAAYNESSLSLSQKAT 540
QY 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLDTITAPNGTYXGNDPXXPXXNMD 600
DB 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLDTITAPNGTYXGNDPXXPXXNMD 600
QY 601 GRNNVENVFINKPOSQTYTIEVQAYNVPGQXFSIAIVN 640
DB 601 GRNNVENVFINKPOSQTYTIEVQAYNVPGQXFSIAIVN 640

RESULT 2

US-10-456-479-4
; Sequence 4, Application US/10456479
; Publication No. US2004007221A1
; GENERAL INFORMATION:
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: TAKIMURA, YASUSHI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: NOMURA, MASAFUMI
; APPLICANT: KOBAYASHI, TOHRU
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 238700USO
; CURRENT APPLICATION NUMBER: US/10/456,479
; CURRENT FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: JP 2002-186387
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: JP 2002-304232
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp. KSM-KP43
US-10-456-479-4

Query Match 99.0%; Score 3029; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 8.7e-261;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 MRXKKKVFSLVLSAAALITVLAALNPSAGAXRXPDLDFKGIQTITDDXXGFSKQOTGAAA 60
DB 1 MRXKKKVFSLVLSAAALITVLAALNPSAGAXRXPDLDFKGIQTITDDXXGFSKQOTGAAA 60
QY 61 FLVSEENVKLLKGLKKLETVANNKHLIXQFNGLIBETKQLEXTGAKILDYIPVAY 120
DB 61 FLVSEENVKLLKGLKKLETVANNKHLIXQFNGLIBETKQLEXTGAKILDYIPVAY 120
QY 121 IVEYEGDVKXSKXXKLEHVESVEPYLPXXIIPDLFTKGA SXIVKXALDPTQXKKEVQLR 180
DB 121 IVEYEGDVKXSKXXKLEHVESVEPYLPXXIIPDLFTKGA SXIVKXALDPTQXKKEVQLR 180
QY 180 IVEYEGDVKXSKXXKLEHVESVEPYLPXXIIPDLFTKGA SXIVKXALDPTQXKKEVQLR 180
DB 180 IVEYEGDVKXSKXXKLEHVESVEPYLPXXIIPDLFTKGA SXIVKXALDPTQXKKEVQLR 180

QY 181 GIEIXIAQXXXNDVYXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGGQGLVAAVADTGLD 240
DB 181 GIEIXIAQXXXNDVYXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGGQGLVAAVADTGLD 240
QY 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGXTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGXTNKGMAPOANLVF 300
QY 301 QSIMDSXGGLGGLPSNLQTLFQOAXSAGARLHTNSWGAUVNGAYTTDSRNVDYVRKNDM 360
DB 301 QSIMDSXGGLGGLPSNLQTLFQOAXSAGARLHTNSWGAUVNGAYTTDSRNVDYVRKNDM 360
QY 361 TILPAAGNEXPNNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRPTKDG 420
DB 361 TILPAAGNEXPNNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRPTKDG 420
QY 421 RIKPDVMAPGTIIISARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
DB 421 RIKPDVMAPGTIIISARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
QY 481 NRGITPKPSLKAALIAAGADIGLGYPNGNGMGRVTLDKSLNVAAYNESSLSLSQKAT 540
DB 481 NRGITPKPSLKAALIAAGADIGLGYPNGNGMGRVTLDKSLNVAAYNESSLSLSQKAT 540
QY 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLDTITAPNGTYXGNDPXXPXXNMD 600
DB 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLDTITAPNGTYXGNDPXXPXXNMD 600
QY 601 GRNNVENVFINKPOSQTYTIEVQAYNVPGQXFSIAIVN 640
DB 601 GRNNVENVFINKPOSQTYTIEVQAYNVPGQXFSIAIVN 640

RESULT 3

US-10-784-870-6
; Sequence 6, Application US/10784870
; Publication No. US20040142837A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/10/784,870
; CURRENT FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US/09/509,814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-10-784-870-6

Query Match 99.0%; Score 3029; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 8.7e-261;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 MRXKKKVFSLVLSAAALITVLAALNPSAGAXRXPDLDFKGIQTITDDXXGFSKQOTGAAA 60
DB 1 MRXKKKVFSLVLSAAALITVLAALNPSAGAXRXPDLDFKGIQTITDDXXGFSKQOTGAAA 60
QY 61 FLVSEENVKLLKGLKKLETVANNKHLIXQFNGLIBETKQLEXTGAKILDYIPVAY 120
DB 61 FLVSEENVKLLKGLKKLETVANNKHLIXQFNGLIBETKQLEXTGAKILDYIPVAY 120

Db 61 FLVESENVKLPKGLQKLETPVANNKLIHQFNGPILBETKQOLEKTKGAKILIDYIPDAY 120
Qy 121 IVEYEGDVASXXXXIHEVESVEPYLPYXIDPOLFTKGASXLVKAALDTKQNKKEVQLR 180
Db 121 IVEYEGDVASATSTIEHVESVEPYLPYRIDPOLFTKGASBLVKAVALDTKQNKKEVQLR 180
Qy 181 GIEIXIQXXXSNDVXYITAKPEYKVMNDVARGIVKADVQOSYGLYGCGQIYVAVDTGLD 240
Db 181 GIEIQIAQFASNDVXYITAKPEYKVMNDVARGIVKADVQOSYGLYGCGQIYVAVDTGLD 240
Qy 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGVAGSVLGNKXTNKGMAPQANLVF 300
Db 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGVAGSVLGNKXTNKGMAPQANLVF 300
Qy 301 QSIIMDSXGGLGGLPSNLQTLFSGQASAGARIHTNSGAAVNGAYTTDSRVNDVYKAKND 360
Db 301 QSIIMDSGGLGGLPSNLQTLFSGQASAGARIHTNSGAAVNGAYTTDSRVNDVYKAKND 360
Qy 361 TILFAAGNEKPNCGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGFTKDG 420
Db 361 TILFAAGNEKPNCGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGFTKDG 420
Qy 421 RIKPDVMAFGTILSARSSLPADSSFWANHDSKYAMGTSMAATPIVAGNVAQLRHFYK 480
Db 421 RIKPDVMAFGTILSARSSLPADSSFWANHDSKYAMGTSMAATPIVAGNVAQLRHFYK 480
Qy 481 NRGITPKPSILKALJAGAADKGLGYPNGNQMGRTVTLKSLNVAVYNESSXLSSTQKAT 540
Db 481 NRGITPKPSILKALJAGAADKGLGYPNGNQMGRTVTLKSLNVAVYNESSXLSSTQKAT 540
Qy 541 YKFTATAGKPLKISLWSDAPASTTASVTLVNDLIVITAPNGTXYVGNDFPKXPKXKND 600
Db 541 YKFTATAGKPLKISLWSDAPASTTASVTLVNDLIVITAPNGTXYVGNDFPKXPKXKND 600
Qy 601 GRNNVENVFINKPQSGTYTIEVOAYNVVPVGPQXFSIAIYN 640
Db 601 GRNNVENVFINKPQSGTYTIEVOAYNVVPVGPQXFSIAIYN 640

RESULT 4
US-10-820-712A-3
Sequence 3, Application US/10820712A
Publication No. US20050026804A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuoshi
APPLICANT: Izawa, Yoshihumi
APPLICANT: Kobayashi, Tohru
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-USO
CURRENT APPLICATION NUMBER: US/10/820, 712A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106708
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp. KSM-KP43
US-10-820-712A-3

Query Match 99.0%; Score 3029; DB 5; Length 640;
Best Local Similarity 93.3%; Pred. No. 8.7e-261;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MKKKKKVFLSVLSAAAILSTVALKNPSAGXARXFDLPDFKGIQTTDDXXGFSKQXGTGAA 60
Db 1 MKKKKKVFLSVLSAAAILSTVALSNPSAGARNFDPDFKGIQTTDDAKGFSKQGTGAA 60

Qy 61 FLVESENVKXXGLXKGLKLETPVANNKLIHQFNGPILBETKQOLEKTKGAKILIDYIPDAY 120
Db 61 FLVESENVKLPKGLQKLETPVANNKLIHQFNGPILBETKQOLEKTKGAKILIDYIPDAY 120
Qy 121 IVEYEGDVASXXXXIHEVESVEPYLPYXIDPOLFTKGASXLVKAALDTKQNKKEVQLR 180
Db 121 IVEYEGDVASATSTIEHVESVEPYLPYRIDPOLFTKGASBLVKAVALDTKQNKKEVQLR 180
Qy 181 GIEIXIQXXXSNDVXYITAKPEYKVMNDVARGIVKADVQOSYGLYGCGQIYVAVDTGLD 240
Db 181 GIEIQIAQFASNDVXYITAKPEYKVMNDVARGIVKADVQOSYGLYGCGQIYVAVDTGLD 240
Qy 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGVAGSVLGNKXTNKGMAPQANLVF 300
Db 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGVAGSVLGNKXTNKGMAPQANLVF 300
Qy 301 QSIIMDSXGGLGGLPSNLQTLFSGQASAGARIHTNSGAAVNGAYTTDSRVNDVYKAKND 360
Db 301 QSIIMDSGGLGGLPSNLQTLFSGQASAGARIHTNSGAAVNGAYTTDSRVNDVYKAKND 360
Qy 361 TILFAAGNEKPNCGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGFTKDG 420
Db 361 TILFAAGNEKPNCGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGFTKDG 420
Qy 421 RIKPDVMAFGTILSARSSLPADSSFWANHDSKYAMGTSMAATPIVAGNVAQLRHFYK 480
Db 421 RIKPDVMAFGTILSARSSLPADSSFWANHDSKYAMGTSMAATPIVAGNVAQLRHFYK 480
Qy 481 NRGITPKPSILKALJAGAADKGLGYPNGNQMGRTVTLKSLNVAVYNESSXLSSTQKAT 540
Db 481 NRGITPKPSILKALJAGAADKGLGYPNGNQMGRTVTLKSLNVAVYNESSXLSSTQKAT 540
Qy 541 YKFTATAGKPLKISLWSDAPASTTASVTLVNDLIVITAPNGTXYVGNDFPKXPKXKND 600
Db 541 YKFTATAGKPLKISLWSDAPASTTASVTLVNDLIVITAPNGTXYVGNDFPKXPKXKND 600
Qy 601 GRNNVENVFINKPQSGTYTIEVOAYNVVPVGPQXFSIAIYN 640
Db 601 GRNNVENVFINKPQSGTYTIEVOAYNVVPVGPQXFSIAIYN 640

RESULT 5
US-10-820-714A-3
Sequence 3, Application US/10820714A
Publication No. US20050214922A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuoshi
APPLICANT: Kobayashi, Tohru
APPLICANT: Sumitomo, Nobuyuki
APPLICANT: Takimura, Yasushi
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251697USO
CURRENT APPLICATION NUMBER: US/10/820, 714A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106709
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp. KSM-KP43
US-10-820-714A-3

Query Match 99.0%; Score 3029; DB 5; Length 640;
Best Local Similarity 93.3%; Pred. No. 8.7e-261;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MKKKKKVFLSVLSAAAILSTVALKNPSAGXARXFDLPDFKGIQTTDDXXGFSKQXGTGAA 60
Db 1 MKKKKKVFLSVLSAAAILSTVALSNPSAGARNFDPDFKGIQTTDDAKGFSKQGTGAA 60

QY 61 FLVESENVKLGKXKGLFTVPANNKLIHXOFNPGILFETKXOIXETGAKIIDYIPDAY 120
DB 61 FLVESENVKLPKGLQKLETPANNKLIHIOFNGLFETKQOLEETGAKIIDYIPDAY 120
QY 121 IVEYEGDVASXXXXIEHVESVEPYLPXXYIDPOLFTKGASXLVKAXALDTKXKXEVOLR 180
DB 121 IVEYEGDVASATSTIEHVESVEPYLPYRIDPOLFTKGASXLVKAXALDTKXKXEVOLR 180
QY 181 GIEIXAOKXXSNDVXYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAVADTGLD 240
DB 181 GIEIXAOKPALSNDVXYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAVADTGLD 240
QY 241 TGRNDSMHEAFRGKITIYALGRTNNANDTNGHGHVAGSVLGNSTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITIYALGRTNNANDTNGHGHVAGSVLGNSTNKGMAPOANLVF 300
QY 301 OSIMDSXGGLGGLPSNLQTLFSGOXSAGARIHTNSKGAAVNGAYTTDSRNVDYVRKNDM 360
DB 301 OSIMDSXGGLGGLPSNLQTLFSGOXSAGARIHTNSKGAAVNGAYTTDSRNVDYVRKNDM 360
QY 361 TILPAAGNEXPGNGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPTKDG 420
DB 361 TILPAAGNEXPGNGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPTKDG 420
QY 421 RIKPDVMAFGTILSARSLSLAPDSSFWANHDSKAYVMGTSMTPTVAGNVAQLREHFVK 480
DB 421 RIKPDVMAFGTILSARSLSLAPDSSFWANHDSKAYVMGTSMTPTVAGNVAQLREHFVK 480
QY 481 NRGITPKPSLLKALILAGADIGLGYPNGNGMGRVTLDKSLNVAVYNESSXLSSTOKAT 540
DB 481 NRGITPKPSLLKALILAGADIGLGYPNGNGMGRVTLDKSLNVAVYNESSXLSSTOKAT 540
QY 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXXPKXXNMD 600
DB 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXXPKXXNMD 600
QY 601 GRNNVENVFINKPOSCTYITIEVOAYNVPGPQXSLAIYN 640
DB 601 GRNNVENVFINKPOSCTYITIEVOAYNVPGPQXSLAIYN 640

RESULT 6
US-09-920-954-8
; Sequence 8, Application US/09920954
; Publication No. US20020064854A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920, 954
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509, 814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-920-954-8

Query Match 99.0%; Score 3028; DB 3; Length 640;
Best Local Similarity 93.3%; Pred. No. 1,1e-260;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 MRXXXXVLSLSAAALISTVALNXPAGKARXEDLPFKGIQTTTDXKXGSKOXQOTGAAA 60
DB 1 MRXXXXVLSLSAAALISTVALNXPAGKARNRDLDFKGIQTTTDXKXGSKOXQOTGAAA 60
QY 61 FLVESENVKLGKXKGLFTVPANNKLIHXOFNPGILFETKXOIXETGAKIIDYIPDAY 120
DB 61 FLVESENVKLPKGLQKLETPANNKLIHIOFNGLFETKQOLEETGAKIIDYIPDAY 120
QY 121 IVEYEGDVASXXXXIEHVESVEPYLPXXYIDPOLFTKGASXLVKAXALDTKXKXEVOLR 180
DB 121 IVEYEGDVASATSTIEHVESVEPYLPYRIDPOLFTKGASXLVKAXALDTKXKXEVOLR 180
QY 181 GIEIXAOKXXSNDVXYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAVADTGLD 240
DB 181 GIEIXAOKPALSNDVXYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAVADTGLD 240
QY 241 TGRNDSMHEAFRGKITIYALGRTNNANDTNGHGHVAGSVLGNSTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITIYALGRTNNANDTNGHGHVAGSVLGNSTNKGMAPOANLVF 300
QY 301 OSIMDSXGGLGGLPSNLQTLFSGOXSAGARIHTNSKGAAVNGAYTTDSRNVDYVRKNDM 360
DB 301 OSIMDSXGGLGGLPSNLQTLFSGOXSAGARIHTNSKGAAVNGAYTTDSRNVDYVRKNDM 360
QY 361 TILPAAGNEXPGNGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPTKDG 420
DB 361 TILPAAGNEXPGNGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPTKDG 420
QY 421 RIKPDVMAFGTILSARSLSLAPDSSFWANHDSKAYVMGTSMTPTVAGNVAQLREHFVK 480
DB 421 RIKPDVMAFGTILSARSLSLAPDSSFWANHDSKAYVMGTSMTPTVAGNVAQLREHFVK 480
QY 481 NRGITPKPSLLKALILAGADIGLGYPNGNGMGRVTLDKSLNVAVYNESSXLSSTOKAT 540
DB 481 NRGITPKPSLLKALILAGADIGLGYPNGNGMGRVTLDKSLNVAVYNESSXLSSTOKAT 540
QY 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXXPKXXNMD 600
DB 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXXPKXXNMD 600
QY 601 GRNNVENVFINKPOSCTYITIEVOAYNVPGPQXSLAIYN 640
DB 601 GRNNVENVFINKPOSCTYITIEVOAYNVPGPQXSLAIYN 640

RESULT 7
US-10-784-870-8
; Sequence 8, Application US/10784870
; Publication No. US20040142837A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/10/784, 870
; CURRENT FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRF
; ORGANISM: Bacillus sp.
US-10-784-870-8

Query Match 99.0%; Score 3028; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 1,1e-260;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
QY 1 MEXKKVFLSVLSAAAILSTVALXNPXAGXARXFDLDFKGIQTTTDXGFSKXQGTGA 60
DB 1 MEXKKVFLSVLSAAAILSTVALXNPXAGXARXFDLDFKGIQTTTDXGFSKXQGTGA 60
QY 61 FVSEENVTLKGLKQLETPANNKGLHIXQNGPILBETKQLEKTKAKILDYIPDYAV 120
DB 61 FVSEENVTLKGLKQLETPANNKGLHIXQNGPILBETKQLEKTKAKILDYIPDYAV 120
QY 121 IVEEGDVXSKXXKXIEHVESVEPYLPXYXIDPOLFTKASXLVKXALDTKQXKEVOLR 180
DB 121 IVEEGDVXSKXXKXIEHVESVEPYLPXYXIDPOLFTKASXLVKXALDTKQXKEVOLR 180
QY 181 GIEKIAQXXXNDVYXITAKPEYKXWMDVARGIVKADVAQSSYGLYGQGIYAVADTGLD 240
DB 181 GIEQIAQPAISNDVYITAKPEYKXWMDVARGIVKADVAQSSYGLYGQGIYAVADTGLD 240
QY 241 TGRNDSMHBAFRGKITLALYALGRTNANDTNGHGTTHVAGSYLGNGXTNKGMAPOANLVF 300
DB 241 TGRNDSMHBAFRGKITLALYALGRTNANDTNGHGTTHVAGSYLGNGXTNKGMAPOANLVF 300
QY 301 GSIMDSXGGLGGLPSNLQTLFSAQXASAGARIHTNSGAAVNGAYTTDSRNVDDYRKNDM 360
DB 301 GSIMDSXGGLGGLPSNLQTLFSAQXASAGARIHTNSGAAVNGAYTTDSRNVDDYRKNDM 360
QY 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENTLPSFGSYADINNHVAQFSRGGTQKG 420
DB 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENTLPSFGSYADINNHVAQFSRGGTQKG 420
QY 421 RIKPVVMAFGTYIISARSSLAPDSSFMANHDSKTYAMGTSMAATPIVAGNVAQLREHFVK 480
DB 421 RIKPVVMAFGTYIISARSSLAPDSSFMANHDSKTYAMGTSMAATPIVAGNVAQLREHFVK 480
QY 481 NRGITPKSSLKAAIAGAADXGLGYPNGNQGMRVTLDKSLNVAAYNESSSLSTSQKAT 540
DB 481 NRGITPKSSLKAAIAGAADXGLGYPNGNQGMRVTLDKSLNVAAYNESSSLSTSQKAT 540
QY 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTXKXVNDPXXPKXXKND 600
DB 541 YSFTATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTXKXVNDPXXPKXXKND 600
QY 601 GRNNVENVFINKPOSGTYYTIEVOAYNVPGVQPSLAIYN 640
DB 601 GRNNVENVFINKPOSGTYYTIEVOAYNVPGVQPSLAIYN 640

RESULT 8
US-09-920-954-4
; Sequence 4, Application US/0920954
; Publication No. US20020064854A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; CURRENT FILING DATE: 2001-06-03
```

PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/J998/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRF
; ORGANISM: Bacillus sp.
US-09-920-954-4

Query Match 98.7%; Score 3020; DB 3; Length 639;
Best Local Similarity 93.4%; Pred. No. 5,5e-260;
Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
QY 4 KKKVFLSVLSAAAILSTVALXNPXAGXARXFDLDFKGIQTTTDXGFSKXQGTGAAPLV 63
DB 3 KKKVFLSVLSAAAILSTVALXNPXAGXARXFDLDFKGIQTTTDXGFSKXQGTGAAPLV 62
QY 64 ESEENVTLKGLKQLETPANNKGLHIXQNGPILBETKQLEKTKAKILDYIPDYAVYE 123
DB 64 ESEENVTLKGLKQLETPANNKGLHIXQNGPILBETKQLEKTKAKILDYIPDYAVYE 122
QY 124 YEGDVXSKXXKXIEHVESVEPYLPXYXIDPOLFTKASXLVKXALDTKQXKEVOLRGIE 183
DB 124 YEGDVOSKRSIEHVESVEPYLPXYXIDPOLFTKASXLVKXALDTKQXKEVOLRGIE 182
QY 184 XIAQXXXNDVYXITAKPEYKXWMDVARGIVKADVAQSSYGLYGQGIYAVADTGLDGR 243
DB 184 EIAQYASNDVYITAKPEYKXWMDVARGIVKADVAQSSYGLYGQGIYAVADTGLDGR 242
QY 244 NDSMHBAFRGKITLALYALGRTNANDTNGHGTTHVAGSYLGNGXTNKGMAPOANLVFOSI 303
DB 243 NDSMHBAFRGKITLALYALGRTNANDTNGHGTTHVAGSYLGNGXTNKGMAPOANLVFOSI 302
QY 304 MDSXGGLGGLPSNLQTLFSAQXASAGARIHTNSGAAVNGAYTTDSRNVDDYRKNDMTL 363
DB 303 MDSXGGLGGLPSNLQTLFSAQXASAGARIHTNSGAAVNGAYTTDSRNVDDYRKNDMTL 362
QY 364 PAAGNEXPNGGTISAPGTAKNAITVGATENTLPSFGSYADINNHVAQFSRGGTQKG 423
DB 363 PAAGNERPNGGTISAPGTAKNAITVGATENTLPSFGSYADINNHVAQFSRGGTQKG 422
QY 424 PDVMAFGTYIISARSSLAPDSSFMANHDSKTYAMGTSMAATPIVAGNVAQLREHFVK 483
DB 423 PDVMAFGTYIISARSSLAPDSSFMANHDSKTYAMGTSMAATPIVAGNVAQLREHFVK 482
QY 484 ITPKSSLKAAIAGAADXGLGYPNGNQGMRVTLDKSLNVAAYNESSSLSTSQKATYX 543
DB 483 ITPKSSLKAAIAGAADXGLGYPNGNQGMRVTLDKSLNVAAYNESSSLSTSQKATYX 542
QY 544 TATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTXKXVNDPXXPKXXKND 603
DB 543 TATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTXKXVNDPXXPKXXKND 602
QY 604 NVENVFINKPOSGTYYTIEVOAYNVPGVQPSLAIYN 640
DB 603 NVENVFINKPOSGTYYTIEVOAYNVPGVQPSLAIYN 639
```

RESULT 9
US-10-784-870-4
; Sequence 4, Application US/10784870
; Publication No. US20040142637A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN


```
; LOCATION: (133)..(133)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (134)..(134)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (147)..(147)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (149)..(149)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (161)..(161)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (166)..(166)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (173)..(173)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (184)..(184)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (188)..(188)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (189)..(189)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (190)..(190)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (195)..(195)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (287)..(287)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (325)..(325)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (370)..(370)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (432)..(432)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (502)..(502)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (532)..(532)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (542)..(542)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (585)..(585)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (592)..(592)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (593)..(593)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (595)..(595)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (596)..(596)
```

```
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (597)..(597)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (612)..(612)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (633)..(633)
; OTHER INFORMATION: Xaa is any amino acid
US-09-920-954-2
```

Query Match 98.6%; Score 3016; DB 3; Length 640;
Best Local Similarity 100.0%; Pred. No. 1,3e-259; Indels 0; Gaps 0;
Matches 640; Conservative 0; Mismatches 0;

```
QY 1 MEXKKKVFSLVLSAAAILSTVALXNPSSAGAXXFDLDFGIQTTDXKXFGSKXQTGA 60
|
|
|
Db 1 MEXKKKVFSLVLSAAAILSTVALXNPSSAGAXXFDLDFGIQTTDXKXFGSKXQTGA 60
|
|
|
QY 61 FLVSEENVKLXKGLKKLETPANNKMLIXQFNGLLEETKQXLEXTGAKIIDYIPDAY 120
|
|
|
Db 61 FLVSEENVKLXKGLKKLETPANNKMLIXQFNGLLEETKQXLEXTGAKIIDYIPDAY 120
|
|
|
QY 121 IVEYEGDVAXSXXXIIEHVESVPEPI.PXYXIDPQLPTKGSXIVKXALDTKQXKEVQJR 180
|
|
|
Db 121 IVEYEGDVAXSXXXIIEHVESVPEPI.PXYXIDPQLPTKGSXIVKXALDTKQXKEVQJR 180
|
|
|
QY 181 GIEIXIAOXXXXNDVYYITAKPEYKVMNDVARGIVRADVQSSYGLYGQGIIVADVDTGLD 240
|
|
|
Db 181 GIEIXIAOXXXXNDVYYITAKPEYKVMNDVARGIVRADVQSSYGLYGQGIIVADVDTGLD 240
|
|
|
QY 241 TGRNDSNMHEAFRGKITLVALGRTNNANDTNGHGHVAGSVLGNXGXTNKGNAPOANLVF 300
|
|
|
Db 241 TGRNDSNMHEAFRGKITLVALGRTNNANDTNGHGHVAGSVLGNXGXTNKGNAPOANLVF 300
|
|
|
QY 301 QSINMSXGGLGGLPSNLQTLFQOAXSAGARITNNSWGAAVNCAYTTDSANVDYVRKNDM 360
|
|
|
Db 301 QSINMSXGGLGGLPSNLQTLFQOAXSAGARITNNSWGAAVNCAYTTDSANVDYVRKNDM 360
|
|
|
QY 361 TILPAAGNEXPMGCTISAPGTAKNAITVGATENLPSFGSYADNINHVAFSSRGPTKDG 420
|
|
|
Db 361 TILPAAGNEXPMGCTISAPGTAKNAITVGATENLPSFGSYADNINHVAFSSRGPTKDG 420
|
|
|
QY 421 RIKPDVMA.PGTXILSARSLAPDSSFMANHDSKTYVMGTSNATPIVAGNVAQLREHFVK 480
|
|
|
Db 421 RIKPDVMA.PGTXILSARSLAPDSSFMANHDSKTYVMGTSNATPIVAGNVAQLREHFVK 480
|
|
|
QY 481 NRGITPKPSLKLKALIAGAADGGLGYPNGNCGWGVTLDKSLNVAVYNESSXILSTSOKAT 540
|
|
|
Db 481 NRGITPKPSLKLKALIAGAADGGLGYPNGNCGWGVTLDKSLNVAVYNESSXILSTSOKAT 540
|
|
|
QY 541 YXFATAGKPLKISLWSDAPASTTASVTLVNDLVLVTAPNGTYGVGDPFXKPKXNDMD 600
|
|
|
Db 541 YXFATAGKPLKISLWSDAPASTTASVTLVNDLVLVTAPNGTYGVGDPFXKPKXNDMD 600
|
|
|
QY 601 GRNNVENYFINXPOSGTYTIRVOAYNVVGPQXPSLAIYN 640
|
|
|
Db 601 GRNNVENYFINXPOSGTYTIRVOAYNVVGPQXPSLAIYN 640
|
|
|
RESULT 11
US-10-784-870-2
; Sequence 2, Application US/10784870
; Publication No. US20040142837A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
```

APPLICANT: NONTURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid

FEATURE:
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:

```
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-10-784-870-2
```

```
Query Match 98.6%; Score 3016; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 1.3e-259;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRXKKRVLSVLSAAIISTVALXNPSAGXAXFPLDPFGIGITTTDXGFSKXOTGAAA 60
DB 1 MRXKKRVLSVLSAAIISTVALXNPSAGXAXFPLDPFGIGITTTDXGFSKXOTGAAA 60
QY 61 FLVESBENVKLGKGLKKTETVPANNKGLHXOPNGPILBETKXKLXTGAKLIDYIPDAY 120
DB 61 FLVESBENVKLGKGLKKTETVPANNKGLHXOPNGPILBETKXKLXTGAKLIDYIPDAY 120
QY 121 IVEYEGDVYKXXXXIEHVESVEPYLPXXYIDPQLFTKGSXLYKXALDTRKXNKEVQLR 180
DB 121 IVEYEGDVYKXXXXIEHVESVEPYLPXXYIDPQLFTKGSXLYKXALDTRKXNKEVQLR 180
QY 181 GLEXIAQXXXXNDVYITIAKPEYKVMNDVARGIVADVAQSSYGLYGQOIIVAVADTGID 240
DB 181 GLEXIAQXXXXNDVYITIAKPEYKVMNDVARGIVADVAQSSYGLYGQOIIVAVADTGID 240
QY 241 TGRNDSMHEAPRGKITLALYALGRTNNANDTNGHGHVAGSYLGNKXTNKGAPQANLVF 300
DB 241 TGRNDSMHEAPRGKITLALYALGRTNNANDTNGHGHVAGSYLGNKXTNKGAPQANLVF 300
QY 301 QSIMDSXGGLGGLPSNLQTLFSGQASAGARIHITNSMGAAVNGAYTTDSRNVDDYVRKXDM 360
DB 301 QSIMDSXGGLGGLPSNLQTLFSGQASAGARIHITNSMGAAVNGAYTTDSRNVDDYVRKXDM 360
```

```
QY 361 TLTPAAGNEXPNGGTTISAPGTAKNAITVGAATENLAPSGSYADNINHVAQPSRGGPTXG 420
DB 361 TLTPAAGNEXPNGGTTISAPGTAKNAITVGAATENLAPSGSYADNINHVAQPSRGGPTXG 420
QY 421 RIKPDVMAPGTIIILARSGLAPDSSFMANHDSKYVMGTSNATPIVAGNVQLREHFK 480
DB 421 RIKPDVMAPGTIIILARSGLAPDSSFMANHDSKYVMGTSNATPIVAGNVQLREHFK 480
QY 481 NFGITPKOSLLKALIIAGADYGLGYPNGOQGWAVTLDDKSLINVAVYNESSXLSTSQKAT 540
DB 481 NFGITPKOSLLKALIIAGADYGLGYPNGOQGWAVTLDDKSLINVAVYNESSXLSTSQKAT 540
QY 541 YXFTATKPKPLKISLWSDAPASTTASVTLVNDLVLTTAPNGTYTVGNDPFXKPKXXNMD 600
DB 541 YXFTATKPKPLKISLWSDAPASTTASVTLVNDLVLTTAPNGTYTVGNDPFXKPKXXNMD 600
QY 601 GRNNVENYFINKPQSGTYTIEVOAYNVVPVGPQXPSLAIYN 640
DB 601 GRNNVENYFINKPQSGTYTIEVOAYNVVPVGPQXPSLAIYN 640
```

```
RESULT 12
US-09-920-954-1
Sequence 1, Application US/09920954
Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKI, KATSUHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIRATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPR: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
```

```
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (501)..(501)
```

```
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
us-09-920-954-1
```

Query Match 98.3%; Score 3007; DB 3; length 639;
Best Local Similarity 100.0%; Pred. No. 7.9e-259;
Matches 637, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

```
4 KKKVFLVLSAAALSTVALXNPSAGXARXPDLDPKIQTITDXXGSKOXOTGAAPFLV 63
3 KKKVFLVLSAAALSTVALXNPSAGXARXPDLPKGIQTITDXXGSKOXOTGAAPFLV 62
64 ESENVKLXKGLXKKLEFVPANNKLIHQFNGPILEETKOKLETKGAILDIYIPYAYIVE 123
63 ESENVKLXKGLXKKLEFVPANNKLIHQFNGPILEETKOKLETKGAILDIYIPYAYIVE 122
124 YEGDVXSKXXKIEHVESVEPYLPYXIDPOLFTKGASXLVKAALDTKXNKREYQLRGIR 183
123 YEGDVXSKXXKIEHVESVEPYLPYXIDPOLFTKGASXLVKAALDTKXNKREYQLRGIR 182
184 XIAQXXSNDVXYTTARDEKVMNDVARGIVKADVAQSSGLVGQGGIVAVADTGLDTR 243
183 XIAQXXSNDVXYTTARDEKVMNDVARGIVKADVAQSSGLVGQGGIVAVADTGLDTR 242
244 NDSMHEAPFGKITLALYALGRTNANDTNGHGVAGSVLGNKXTNGMAPLANLVFQSI 303
243 NDSMHEAPFGKITLALYALGRTNANDTNGHGVAGSVLGNKXTNGMAPLANLVFQSI 302
304 MDSYGLGGLPSNIQTLFQQAQXAGARIHTNSGAAVNGAYTTDSRVDDYVRKNDMTIL 363
303 MDSYGLGGLPSNIQTLFQQAQXAGARIHTNSGAAVNGAYTTDSRVDDYVRKNDMTIL 362
364 PAAGNEXPNGGTISAPETAKNAITTVGATEMLRSPFSYADNINHVQOFSSRGPTKGRIR 423
363 PAAGNEXPNGGTISAPETAKNAITTVGATEMLRSPFSYADNINHVQOFSSRGPTKGRIR 422
424 PDVWAPGTXTILSARSSILAPDSSFMANHDSKYAVNGTSMATPIVAGVNAQLRHFVNRG 483
423 PDVWAPGTXTILSARSSILAPDSSFMANHDSKYAVNGTSMATPIVAGVNAQLRHFVNRG 482
484 ITPKPSILKAAALTAGAADXGLGYPNGNGGGRVTLTDKSLNVAVNBSSXLSTSQKATYXF 543
483 ITPKPSILKAAALTAGAADXGLGYPNGNGGGRVTLTDKSLNVAVNBSSXLSTSQKATYXF 542
```


Qy	544	TATACKPLKSLVWSDAPASTASTATLVNDLDTTAPRGTYVNGDPKPKXXNWDGN	603	OTHER INFORMATION: Xaa is any amino acid
Db	543	TATACKPLKSLVWSDAPASTASTATLVNDLDTTAPRGTYVNGDPKPKXXNWDGN	602	FEATURE:
Qy	604	NVENVFINKPQSGTTITIEVQAYNVPPVGPQXPSLAIYN	640	NAME/KEY: misc feature
Db	603	NVENVFINKPQSGTTITIEVQAYNVPPVGPQXPSLAIYN	639	LOCATION: (102)..(102)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (105)..(105)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (128)..(128)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (130)..(130)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (131)..(131)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (132)..(132)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (133)..(133)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (146)..(146)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (148)..(148)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (160)..(160)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (165)..(165)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (172)..(172)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (183)..(183)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (187)..(187)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (188)..(188)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (189)..(189)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (194)..(194)
				OTHER INFORMATION: Xaa is any amino acid
				FEATURE:
				NAME/KEY: misc feature
				LOCATION: (286)..(286)
				OTHER INFORMATION: Xaa is any amino acid

FEATURE:
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-10-784-870-1

Query Match 98.3%; Score 3007; DB 4; Length 639;
Best Local Similarity 100.0%; Pred. NO. 7.9e-259;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KKKVFLSVLSAAALSTVALKNPSAGXAXFDLPFGKIQTITDDXGFSKXQGTGAAPLV 63
DB 3 KKKVFLSVLSAAALSTVALKNPSAGXAXFDLPFGKIQTITDDXGFSKXQGTGAAPLV 62
QY 64 ESEVVKLKKGLKXKLETVPANNKLIHXQFNGPILEETXQILETGAKTLDYIPYAYIVE 123
DB 63 ESEVVKLKKGLKXKLETVPANNKLIHXQFNGPILEETXQILETGAKTLDYIPYAYIVE 122

QY 124 YEGDVXSKXXXIEHVESVEPLPYXIDPOLFTKGASXIVKAXALDTKONKKEVQLAGIE 183
DB 123 YEGDVXSKXXXIEHVESVEPLPYXIDPOLFTKGASXIVKAXALDTKONKKEVQLAGIE 182
QY 184 XIAQXXXNDVXYITTAPEYKVMNDVARGIVKADVAOSYGLYGQGIIVAAVDTGLDTGR 243
DB 183 XIAQXXXNDVXYITTAPEYKVMNDVARGIVKADVAOSYGLYGQGIIVAAVDTGLDTGR 242
QY 244 NDSMHEAFRGKITALVALGRTNANDTNGHTVAGSVLGNKTKGMAPQANLVFQSI 303
DB 243 NDSMHEAFRGKITALVALGRTNANDTNGHTVAGSVLGNKTKGMAPQANLVFQSI 302
QY 304 MDSXGIGLGLPSNQLTLPQXASAGARITHNSWGAANGAYATDSRVNDVVRNDMTIL 363
DB 303 MDSXGIGLGLPSNQLTLPQXASAGARITHNSWGAANGAYATDSRVNDVVRNDMTIL 362
QY 364 PAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININVAOFSSRGPTDGR 423
DB 363 PAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININVAOFSSRGPTDGR 422
QY 424 PDVAPGTXTILSARSSILAPDSSFVANHDSKYAYWGTSMATPIVAGNVAQLREHFVXRG 483
DB 423 PDVAPGTXTILSARSSILAPDSSFVANHDSKYAYWGTSMATPIVAGNVAQLREHFVXRG 482
QY 484 ITPKPSLLKALITGAADXGLGYPNGOGNGRVTLDLSLVAAVYNESSXISTGKATYXP 543
DB 483 ITPKPSLLKALITGAADXGLGYPNGOGNGRVTLDLSLVAAVYNESSXISTGKATYXP 542
QY 544 TATAGKPLKISLVWSDAPASTASVTLVNDLVTAPNGTXVYVNDPXXPXKNMGRN 603
DB 543 TATAGKPLKISLVWSDAPASTASVTLVNDLVTAPNGTXVYVNDPXXPXKNMGRN 602
QY 604 NVENVFINXPOSGTYTTEVOAIVNVPGPQXFSIAIVN 640
DB 603 NVENVFINXPOSGTYTTEVOAIVNVPGPQXFSIAIVN 639

RESULT 14
US-09-985-689A-1
Sequence 1, Application US/09985689A
Publication No. US2003002351A1
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKETI, KATSUHISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483US0
CURRENT APPLICATION NUMBER: US/09/985,689A
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-1

Query Match 70.4%; Score 2155; DB 3; Length 434;
Best Local Similarity 96.3%; Pred. NO. 4e-183;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 207 NDVARGIVKADVAOSYGLYGQGIIVAAVDTGLDTGRNDSMHEAFRGKITALVALGRTN 266

Db 1 NDVARGIVADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHBAFRGKITALVALGRTN 60
Qy 267 NANDNNGHGTHTVAGSTLGNKXTTKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFESQAXS 326
Db 61 NANDNNGHGTHTVAGSVLGNSTTKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFESQAXS 120
Qy 327 AGARITHNSMGAANVAGATTTDSRNVDDYRKNDMTLLFAAGNEXPNGGTISAPGTAKNAI 386
Db 121 AGARITHNSMGAANVAGATTTDSRNVDDYRKNDMTLLFAAGNEXPNGGTISAPGTAKNAI 180
Qy 387 TVGATENLRPSFGSYADNINHVAFSSRGPPTDGRIKPDVMAPGTYILSARSLAPDSF 446
Db 181 TVGATENLRPSFGSYADNINHVAFSSRGPPTDGRIKPDVMAPGTYILSARSLAPDSF 240
Qy 447 WANHSKAYMGSTMAATPIVAGNVAQLREHFVKRGTIPKPSLKAALIAAGADVGLEY 506
Db 241 WANHSKAYMGSTMAATPIVAGNVAQLREHFVKRGTIPKPSLKAALIAAGADVGLEY 300
Qy 507 PNGOQMGKRVTLDKSLNVAAYNNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 566
Db 301 PNGOQMGKRVTLDKSLNVAAYNNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 360
Qy 567 SVTLVNDLDELVTAPNGTXYVGNDFXKPKXXNMDGRNNEVFINXPOSGTYTIEVOAYN 626
Db 361 SVTLVNDLDELVTAPNGTXYVGNDFXKPKXXNMDGRNNEVFINXPOSGTYTIEVOAYN 420
Qy 627 VPVGPOXFSLAIYN 640
Db 421 VPVGPOXFSLAIYN 434

RESULT 15

US-09-985-689A-2
; Sequence 2, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-2

Query Match 70.4%; Score 2155; DB 3; Length 434;
Best Local Similarity 96.3%; Pred. No. 4e-183;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 207 NDVARGIVADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHBAFRGKITALVALGRTN 266
Db 1 NDVARGIVADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHBAFRGKITALVALGRTN 60
Qy 267 NANDNNGHGTHTVAGSVLGNSTTKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFESQAXS 326
Db 61 NANDNNGHGTHTVAGSVLGNSTTKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFESQAXS 120
Qy 327 AGARITHNSMGAANVAGATTTDSRNVDDYRKNDMTLLFAAGNEXPNGGTISAPGTAKNAI 386

Db 121 AGARITHNSMGAANVAGATTTDSRNVDDYRKNDMTLLFAAGNEXPNGGTISAPGTAKNAI 180
Qy 387 TVGATENLRPSFGSYADNINHVAFSSRGPPTDGRIKPDVMAPGTYILSARSLAPDSF 446
Db 181 TVGATENLRPSFGSYADNINHVAFSSRGPPTDGRIKPDVMAPGTYILSARSLAPDSF 240
Qy 447 WANHSKAYMGSTMAATPIVAGNVAQLREHFVKRGTIPKPSLKAALIAAGADVGLEY 506
Db 241 WANHSKAYMGSTMAATPIVAGNVAQLREHFVKRGTIPKPSLKAALIAAGADVGLEY 300
Qy 507 PNGOQMGKRVTLDKSLNVAAYNNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 566
Db 301 PNGOQMGKRVTLDKSLNVAAYNNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 360
Qy 567 SVTLVNDLDELVTAPNGTXYVGNDFXKPKXXNMDGRNNEVFINXPOSGTYTIEVOAYN 626
Db 361 SVTLVNDLDELVTAPNGTXYVGNDFXKPKXXNMDGRNNEVFINXPOSGTYTIEVOAYN 420
Qy 627 VPVGPOXFSLAIYN 640
Db 421 VPVGPOXFSLAIYN 434

Search completed: April 7, 2006, 18:03:41
Job time : 119.593 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 17:32:32 ; Search time 138.608 Seconds
(without alignments)
2028.756 Million cell updates/sec

Title: US-10-784-870-2
Perfect score: 3059
Sequence: 1 MRXXXXVFLSLASAAIIST.....EVQAVNPVGPQXPSLAIYN 640

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*
9: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3029	99.0	640	AA117090	Aay17090 Bacillus
2	3029	99.0	640	ADM40773	Adm40773 Alkaline
3	3029	99.0	640	ADSI4437	Adsi4437 Bacillus
4	3029	99.0	641	ADT49613	Adt49613 Bacillus
5	3028	99.0	640	AA117091	Aay17091 Bacillus
6	3020	98.7	639	AA117089	Aay17089 Bacillus
7	3016	98.6	640	AA117088	Aay17088 An alkali
8	3007	98.3	639	AA117087	Aay17087 An alkali
9	2722.5	89.0	641	AAW89547	AAW89547 Bacillus
10	2414	78.9	636	AAW89548	AAW89548 Bacillus
11	2155	70.4	434	AAW50080	AAW50080 Bacillus
12	2155	70.4	434	AAW50081	AAW50081 Bacillus
13	2155	70.4	434	ADY33778	Ady33778 Bacillus
14	2155	70.4	434	ADZ51758	Adz51758 Mutant Ba
15	2155	70.4	434	ADZ51757	Adz51757 Wild-type
16	2155	70.4	434	ADL25802	Adl25802 Bacillus
17	2155	70.4	434	ADM40779	Adm40779 Alkaline
18	2155	70.4	434	ADM40771	Adm40771 Mature al
19	2155	70.4	434	ADM40780	Adm40780 Alkaline
20	2155	70.4	434	ADSI4427	Adsi4427 Bacillus
21	2155	70.4	434	ADT49604	Adt49604 Bacillus
22	2150	70.3	434	ADSI4441	Adsi4441 Bacillus
23	2147	70.2	434	ADSI4438	Adsi4438 Bacillus
24	2147	70.2	434	ADSI4442	Adsi4442 Bacillus

25	2146	70.2	434	8	ADSI4439	Adsi4439 Bacillus
26	2140	70.0	434	8	ADSI4440	Adsi4440 Bacillus
27	2138	69.9	434	8	ADSI4443	Adsi4443 Bacillus
28	2133	69.7	434	8	ADSI4444	Adsi4444 Bacillus
29	2118	69.2	434	8	ADSI4445	Adsi4445 Alkaline
30	2103	68.7	436	8	ADM40787	Adm40787 Bacillus
31	2082	68.1	434	5	AAW50085	AAW50085 Bacillus
32	2082	68.1	434	7	ADZ51762	Adz51762 Mutant Ba
33	2082	68.1	434	8	ADM40784	Adm40784 Alkaline
34	2064.5	67.5	433	8	ADSI52082	Adsi52082 Bacillus
35	2061.5	67.4	433	8	ADSI52083	Adsi52083 Bacillus
36	2060.5	67.4	433	5	AAW50086	AAW50086 Bacillus
37	2060.5	67.4	433	7	ADZ51763	Adz51763 Mutant Ba
38	2060.5	67.4	433	8	ADM40785	Adm40785 Alkaline
39	2060.5	67.4	433	8	ADSI52010	Adsi52010 Bacillus
40	2060.5	67.4	433	8	ADSI52016	Adsi52016 Bacillus
41	2060.5	67.4	433	8	ADSI52011	Adsi52011 Bacillus
42	2060.5	67.4	433	8	ADSI52001	Adsi52001 Bacillus
43	2059.5	67.3	433	8	ADSI52062	Adsi52062 Bacillus
44	2059.5	67.3	433	8	ADSI52054	Adsi52054 Bacillus
45	2058.5	67.3	433	8	ADSI52054	Adsi52054 Bacillus

ALIGNMENTS

RESULT 1
AA117090
ID AAY17090 strand; protein, 640 AA.
XX
AC AAY17090;
XX
DT 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
DE Bacillus alkaline protease.
XX
KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW washing composition; oxidizing agent.
XX
OS Bacillus sp.
XX
PN WO9918218-A1.
XX
PD 15-APR-1999.
XX
PF 07-OCT-1998; 98WO-JP004528.
XX
PR 07-OCT-1997; 97JP-00274570.
XX
PA (KAOS) KAO CORP.
XX
PI Takaiwa M, Okuda M, SaeKI K, Kubota H, Hitomi J, Kageyama Y,
PI Shikata S, Nomura M,
DR WPI: 1999-287736/27.
DR N-PSDB; AAX37278.
PT
XX Alkali protease from Bacillus used in washing powders.
XX
PS Disclosure: Page 58-63; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
XX Bacillus. The proteases ability to digest casein is not inhibited by
XX oleic acid and they have a high stability to oxidizing agents. The
XX alkaline protease of the invention has the following properties: (a) It
XX is active over the pH range 4-13 and has at least 80% of its optimum
XX activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
XX stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
XX its ability to digest casein is not inhibited by oleic acid; (e) it has
XX molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
XX used as enzymes in washing compositions for use in automatic dishwashers
XX and for washing clothes. The stability to oxidizing agents allows the

CC enzyme to be an effective component of washing compositions including
CC bleachers. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)

XX Sequence 640 AA;

Query Match 99.0%; Score 3029; DB 2; Length 640;

Best Local Similarity 93.3%; Pred. No. 1.6e-237; Indels 0; Gaps 0;

Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
OY 1 MRKKKKVFLSVLSAAAILSTVALXNPSAGAXRFDLPFKGIQTITDXXGFSKQGTGAAA 60
DB 1 MRKKKKVFLSVLSAAAILSTVALSNPSAGARNFPLDFKGIQTITDAGFSKQGTGAAA 60
OY 61 FLVSEENVKLGKGLKXKLETPANNKLIHQFNPILEETKQJLEXTGAKILDIYIPYAY 120
DB 61 FLVSEENVKLGKGLKXKLETPANNKLIHQFNPILEETKQJLEXTGAKILDIYIPYAY 120
OY 121 IVEYEGDVXSKXXXIEHVESVEPYLPYXXIDPQLFTKASXLVKXALDTKQXKEVQLR 180
DB 121 IVEYEGDVXSKATSTIEHVESVEPYLPYRIDPQLFTKASBLVKAVLDTKQKKEVQLR 180
OY 181 GIEKIAQXXXSNDVYITAKPEYKVMNDVARGIVADVAQSSYGLYGGQIVAAVDGLD 240
DB 181 GIEQIAQFAISNDVLYITAKPEYKVMNDVARGIVADVAQSSYGLYGGQIVAAVDGLD 240
OY 241 TGRNDSMHEAFRGKITLALYALGRNNANDTNGHGTIVAGSYLGNXTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRNNANDTNGHGTIVAGSYLGNXTNKGMAPOANLVF 300
OY 301 OSIIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSWGAANVGAYTTDSRVDDYVRKNDM 360
DB 301 OSIIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSWGAANVGAYTTDSRVDDYVRKNDM 360
OY 361 TILFAAGNEXPNNGTISAPGTAKNAITVGATENLRPSFGSYADININVAQSSRGPTKDG 420
DB 361 TILFAAGNEXPNNGTISAPGTAKNAITVGATENLRPSFGSYADININVAQSSRGPTKDG 420
OY 421 RIKPDVWAPGTXYILSARSSSLAPDSSFANNDSKYAYNGGTSMAPIYVAGNVAQLREHFVK 480
DB 421 RIKPDVWAPGTXYILSARSSSLAPDSSFANNDSKYAYNGGTSMAPIYVAGNVAQLREHFVK 480
OY 481 NRGITPKPSLLKALIIAGAADIGLGYPNGNGMGWRTLDKSLNVAAYNESSSLSTSQKAT 540
DB 481 NRGITPKPSLLKALIIAGAADIGLGYPNGNGMGWRTLDKSLNVAAYNESSSLSTSQKAT 540
OY 541 YKFTATACKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTYVGNDFEXPPXXKND 600
DB 541 YKFTATACKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTYVGNDFEXPPXXKND 600
OY 601 GRNNVENVFINXPOSGETTIEVOAYNVVPGQPSLAIVN 640
DB 601 GRNNVENVFINAPOSGETTIEVOAYNVVPGQPSLAIVN 640
```

RESULT 2

ADM40773 standard; protein; 640 AA.

XX ADM40773;

XX 01-JUN-2004 (first entry)

XX Alkaline protease from *Bacillus* sp. KSM-KP43.

XX alkaline protease; laundry detergent; bleaching agent; detergent;

XX denture-cleaning agent; enzyme.

XX *Bacillus* sp.; KSM-KP43.

XX US2004072321-A1.

XX 15-APR-2004.

XX 09-JUN-2003; 2003US-00456479.
XX 26-JUN-2002; 2002JP-00186387.
XX 18-OCT-2002; 2002JP-00304232.

XX (KAO) KAO CORP.

XX Sato T, Okuda M, Takimura Y, Sumitomo N, Nomura M, Kobayashi T;

XX MPI; 2004-328572/30.

XX N-PSDB; ADM40772.

PT New alkaline protease having a mutant prepro sequence where amino acid
PT residues at positions 52, 75 and 142 are substituted with another amino
PT acid residue, useful as enzyme component of laundry detergents, or
PT bleaching agents.

XX Disclousure; SEQ ID NO 4; 29pp; English.

CC The invention relates to an alkaline protease having a prepro sequence.

CC The prepro sequence is a mutant sequence of SEQ ID NO: 1, or an amino
CC acid sequence having 80% homology or higher to SEQ ID NO: 1, where amino
CC acid residues at: (a) position 52 is substituted by aspartic acid or
CC arginine; (b) position 75 is substituted by alanine or arginine; and (c)
CC position 142 is substituted by lysine. The alkaline protease is useful as
CC an enzyme which can be incorporated into laundry detergents, bleaching
CC agents, detergents for cleaning hard surfaces or drainpipes, denture-
CC cleaning agents, and detergents for sterilizing medical apparatus. The
CC present sequence represents alkaline protease from *Bacillus* sp. KSM-KP43.

XX Sequence 640 AA;

Query Match 99.0%; Score 3029; DB 8; Length 640;

Best Local Similarity 93.3%; Pred. No. 1.6e-237; Indels 0; Gaps 0;

Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
OY 1 MRKKKKVFLSVLSAAAILSTVALXNPSAGAXRFDLPFKGIQTITDXXGFSKQGTGAAA 60
DB 1 MRKKKKVFLSVLSAAAILSTVALSNPSAGARNFPLDFKGIQTITDAGFSKQGTGAAA 60
OY 61 FLVSEENVKLGKGLKXKLETPANNKLIHQFNPILEETKQJLEXTGAKILDIYIPYAY 120
DB 61 FLVSEENVKLGKGLKXKLETPANNKLIHQFNPILEETKQJLEXTGAKILDIYIPYAY 120
OY 121 IVEYEGDVXSKXXXIEHVESVEPYLPYXXIDPQLFTKASXLVKXALDTKQXKEVQLR 180
DB 121 IVEYEGDVXSKATSTIEHVESVEPYLPYRIDPQLFTKASBLVKAVLDTKQKKEVQLR 180
OY 181 GIEKIAQXXXSNDVYITAKPEYKVMNDVARGIVADVAQSSYGLYGGQIVAAVDGLD 240
DB 181 GIEQIAQFAISNDVLYITAKPEYKVMNDVARGIVADVAQSSYGLYGGQIVAAVDGLD 240
OY 241 TGRNDSMHEAFRGKITLALYALGRNNANDTNGHGTIVAGSYLGNXTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRNNANDTNGHGTIVAGSYLGNXTNKGMAPOANLVF 300
OY 301 OSIIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSWGAANVGAYTTDSRVDDYVRKNDM 360
DB 301 OSIIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSWGAANVGAYTTDSRVDDYVRKNDM 360
OY 361 TILFAAGNEXPNNGTISAPGTAKNAITVGATENLRPSFGSYADININVAQSSRGPTKDG 420
DB 361 TILFAAGNEXPNNGTISAPGTAKNAITVGATENLRPSFGSYADININVAQSSRGPTKDG 420
OY 421 RIKPDVWAPGTXYILSARSSSLAPDSSFANNDSKYAYNGGTSMAPIYVAGNVAQLREHFVK 480
DB 421 RIKPDVWAPGTXYILSARSSSLAPDSSFANNDSKYAYNGGTSMAPIYVAGNVAQLREHFVK 480
OY 481 NRGITPKPSLLKALIIAGAADIGLGYPNGNGMGWRTLDKSLNVAAYNESSSLSTSQKAT 540
DB 481 NRGITPKPSLLKALIIAGAADIGLGYPNGNGMGWRTLDKSLNVAAYNESSSLSTSQKAT 540
```

Qy 541 YKFTATAGKPLKLSLVWSDAPASTTASVTLVNDLDTVITAPNGTXYVGNPFXXPKXXNMD 600
 Db 541 YKFTATAGKPLKLSLVWSDAPASTTASVTLVNDLDTVITAPNGTXYVGNPFXXPKXXNMD 600
 Qy 601 GRNNVENVPINXPOSGTYYTIEVOAYNVPGPOXFSALAYN 640
 Db 601 GRNNVENVPINXPOSGTYYTIEVOAYNVPGPOXFSALAYN 640

RESULT 3

ADSI4437
 ID ADSI4437 standard; protein; 640 AA.

XX ADSI4437;

XX 30-DEC-2004 (first entry)

XX Bacillus alkaline protease KP43.

XX protease; enzyme; alkaline protease; laundry detergent; KP43.

XX Bacillus sp.; KSM-KP43.

XX Key Location/Qualifiers

XX Peptide 1..206

XX Protein 207..640

XX EPI466962-A1.

XX 13-OCT-2004.

XX 08-APR-2004; 2004EP-00008604.

XX 10-APR-2003; 2003JP-00106709.

XX (KAOS) KAO CORP.

XX Okuda M, Sato T, Takimura Y, Sumitomo N, Kobayashi T;

XX WPI, 2004-711313/70.

XX N-PSDB; ADSI4428.

XX New engineered alkaline protease, useful particularly in laundry

XX detergents, comprising specified amino acids at particular positions.

XX Disclosure; SEQ ID NO 1; 31pp; English.

XX The invention relates to a novel alkaline protease. The new alkaline

XX protease comprises an amino acid sequence in which one or more amino acid

XX ADSI4427, or at positions corresponding to these positions are: position

XX 15 (thiethidine), position 16 (threonine or glutamine), position 166

XX (glycine), position 167 (valine), position 187 (serine), position 346

XX (arginine), and position 405 (aspartic acid). The alkaline protease is

XX useful in industry particularly in laundry detergents, but also e.g. in

XX fibre modifying agents, leather processing agents, cosmetic compositions,

XX bath additives, food-modifying agents, and pharmaceuticals. The present

XX sequence represents the wild-type Bacillus sp. KSM-KP43 alkaline

XX protease, KP43. The sequence is shown in the sequence listing as part of

XX SEQ ID NO:1

XX Sequence 640 AA;

XX Query Match 99.0%; Score 3029; DB 8; Length 640;

XX Best Local Similarity 93.3%; Pred. No. 1.6e-237;

XX Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MKKKKKVPLSTVSAALISTVALNPSAKAXPDLDFGIGITTTDXKFSKQXQTGA 60

Db 1 MKKKKKVPLSTVSAALISTVALNPSAKAXPDLDFGIGITTTDXKFSKQXQTGA 60

Qy 61 FLVSEENVGLKGLKLETVPANNGLHIQFNGPILLETQKXLEXTGAKLIDYIPVAY 120

Db 61 FLVSEENVGLKGLKLETVPANNGLHIQFNGPILLETQKXLEXTGAKLIDYIPVAY 120
 Qy 121 IYVEEGDVXSKXXXIEHVESVEBYLPXYIIDPOLFTKGASXLVKANALDTKQXNKEVOLR 180
 Db 121 IYVEEGDVXSKXXXIEHVESVEBYLPXYIIDPOLFTKGASXLVKANALDTKQXNKEVOLR 180
 Qy 181 GIEIXIAXXXXSNDVYXITAKPEYKMNVDVARGIVKADVQASVGLYGQGIYVAADPTGLD 240
 Db 181 GIEIXIAXXXXSNDVYXITAKPEYKMNVDVARGIVKADVQASVGLYGQGIYVAADPTGLD 240
 Qy 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGTGVASVYGNKXTNKGMAPOANLVP 300
 Db 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGTGVASVYGNKXTNKGMAPOANLVP 300
 Qy 301 QSIMDSXGGLGGLPSNLQTLFSGQASGARIHTNSGAAVNGAYTTDSRVDDYVRKNDM 360
 Db 301 QSIMDSXGGLGGLPSNLQTLFSGQASGARIHTNSGAAVNGAYTTDSRVDDYVRKNDM 360
 Qy 361 TLIFPAAGNEXPNCGTISAPGTAKNATVGAATENTLRPSFGSYADNINRVAQFSSRGPTKOG 420
 Db 361 TLIFPAAGNEXPNCGTISAPGTAKNATVGAATENTLRPSFGSYADNINRVAQFSSRGPTKOG 420
 Qy 421 RIKPDVMAFGTYILSARSSLAPDSFMANHDSKYAVMGTSMATPIVAGVNAQLRSHFYK 480
 Db 421 RIKPDVMAFGTYILSARSSLAPDSFMANHDSKYAVMGTSMATPIVAGVNAQLRSHFYK 480
 Qy 481 NRGITPKPSLLKALIALAGADYGLGYPNGNGMGRTYLDKSLNVAAYNNESSXLSTSQKAT 540
 Db 481 NRGITPKPSLLKALIALAGADYGLGYPNGNGMGRTYLDKSLNVAAYNNESSXLSTSQKAT 540
 Qy 541 YKFTATAGKPLKLSLVWSDAPASTTASVTLVNDLDTVITAPNGTXYVGNPFXXPKXXNMD 600
 Db 541 YKFTATAGKPLKLSLVWSDAPASTTASVTLVNDLDTVITAPNGTXYVGNPFXXPKXXNMD 600
 Qy 601 GRNNVENVPINXPOSGTYYTIEVOAYNVPGPOXFSALAYN 640
 Db 601 GRNNVENVPINXPOSGTYYTIEVOAYNVPGPOXFSALAYN 640

RESULT 4

ID ADT49613 standard; protein; 641 AA.

XX ADT49613;

XX 30-DEC-2004 (first entry)

XX Bacillus alkaline protease.

XX Alkaline protease; detergent; fiber modification; leather processing;

XX cosmetic; bath additives; food-modification; pharmaceutical; enzyme.

XX Bacillus sp. KSM-KP43.

XX Key Location/Qualifiers

XX Peptide 1..206

XX Protein 207..641

XX /note="specifically claimed mature protein (SEQ ID 1)"

XX EPI466970-A1.

XX 13-OCT-2004.

XX 08-APR-2004; 2004EP-00008605.

XX 10-APR-2003; 2003JP-00106708.

XX (KAOS) KAO CORP.

XX Sato T, Okuda M, Koyama S, Izawa Y, Kobayashi T;

XX WPI, 2004-711317/70.

DR N-PSDB; ADT94605.

XX New engineered alkaline protease with improved activity and thermal
PT stability, useful particularly in detergents such as laundry detergents.

XX Discloure; Page 19-25; 35pp; English.

XX The invention relates to an alkaline protease and its encoding gene. The
CC alkaline protease can be expressed by standard recombinant methodology.
CC The alkaline protease is useful in the industry particularly in
CC detergents such as laundry detergents, but also in fiber modifying
CC agents, leather processing agents, cosmetic compositions, bath additives,
CC food-modifying agents, and pharmaceuticals. The enzyme has good activity
CC and thermal stability. The present sequence represents an alkaline
CC protease from *Bacillus* sp. KSM-KP43.

XX Sequence 641 AA;

Query Match 99.0%; Score 3029; DB 8; Length 641;

Best Local Similarity 93.3%; Pred. No. 1.6e-237;

Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```

QY 1 MRKKKKVFLSVLSAAAILSTVALKNPSAGAXKXFDLDFKGIQTITDDXKGFSGKQGTGAAA 60
DB 1 MRKKKKVFLSVLSAAAILSTVALSNPSAGARNPDLDFKGIQTITDAGKFGSQGTGAAA 60
QY 61 FLVSEENVKLGKGLKQLETPANNKRLHXQFNGLPILFTKQXLEXTGAKLIDYIPDYAX 120
DB 61 FLVSEENVKLPKGLQKLETPANNKRLHXQFNGLPILFTKQXLEXTGAKLIDYIPDYAX 120
QY 121 IVEYEGDVKSXXXXIEHVESVEPYLPYXXIDPOLFTKGASXLVKAALDTKQNKXEVQLR 180
DB 121 IVEYEGDVKSANSTIEHVESVEPYLPYRIDPOLFTKGASXLVKAALDTKQNKXEVQLR 180
QY 181 GIEKIAQXXXXNDVYITAKPEYKVMNDVARGIVADVAOSSYGLYGGQGIYAVALDTGLD 240
DB 181 GIEKIAQFAISNDVYITAKPEYKVMNDVARGIVADVAOSSYGLYGGQGIYAVALDTGLD 240
QY 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGHVAGSVLGNSTNGKGAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGHVAGSVLGNSTNGKGAPOANLVF 300
QY 301 QSIMDSXGGLGGLPSNLQTLFSGQASAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDM 360
DB 301 QSIMDSGGGLGGLPSNLQTLFSGQAYSAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDM 360
QY 361 TILFPAAGNEXNGGTTISAPGTAKNAITVGATENLPSFGSYADININHYAQSRSRGTYDG 420
DB 361 TILFPAAGNEXNGGTTISAPGTAKNAITVGATENLPSFGSYADININHYAQSRSRGTYDG 420
QY 421 RIKPVMAPGTTILSARSSILAPDSSPMANHDSKYVMGTSMAFTIVAGNVAKLREHFVK 480
DB 421 RIKPVMAPGTTILSARSSILAPDSSPMANHDSKYVMGTSMAFTIVAGNVAKLREHFVK 480
QY 481 NRGITPKPSLKAALIAAGADYGLYPNGOQMGHVTLDKSLNVAVYNESSLSTSQKAT 540
DB 481 NRGITPKPSLKAALIAAGADYGLYPNGOQMGHVTLDKSLNVAVYNESSLSTSQKAT 540
QY 541 YXFTNATAGPKLISLVMSDAPASTTASVTLVNDLVLVTAPNGYIYVANDFPKXXKND 600
DB 541 YSFTATAGPKLISLVMSDAPASTTASVTLVNDLVLVTAPNGYIYVANDFPKXXKND 600
QY 601 GRNNVENVFINKPOSGYTTIEVOANVVPQXPSLAIYN 640
DB 601 GRNNVENVFINKPOSGYTTIEVOANVVPQXPSLAIYN 640

```

RESULT 5
AA117091

XX AA117091 standard: protein; 640 AA.

AC AA117091;

XX

DT 20-MAR-2003 (revised)

DT 21-JUL-1999 (first entry)

DB Bacillus alkaline protease.

KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;

XX washing composition; oxidizing agent.

OS Bacillus sp.

XX W09918218-A1.

XX 15-APR-1999.

XX 07-OCT-1998; 98WO-JP004528.

XX 07-OCT-1997; 97JP-00274570.

XX (KAOS) KAO CORP.

PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;

XX Shikata S, Nomura M;

XX WPI; 1999-287736/27.

DR N-PSDB; AAX37279.

XX Alkali protease from Bacillus used in washing powders.

XX Discloure; Page 63-68; 71pp; Japanese.

CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; c) Its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)

XX Sequence 640 AA;

Query Match 99.0%; Score 3028; DB 2; Length 640;

Best Local Similarity 93.3%; Pred. No. 1.9e-237;

Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```

QY 1 MRKKKKVFLSVLSAAAILSTVALKNPSAGAXKXFDLDFKGIQTITDDXKGFSGKQGTGAAA 60
DB 1 MRKKKKVFLSVLSAAAILSTVALSNPSAGARNPDLDFKGIQTITDAGKFGSQGTGAAA 60
QY 61 FLVSEENVKLGKGLKQLETPANNKRLHXQFNGLPILFTKQXLEXTGAKLIDYIPDYAX 120
DB 61 FLVSEENVKLPKGLQKLETPANNKRLHXQFNGLPILFTKQXLEXTGAKLIDYIPDYAX 120
QY 121 IVEYEGDVKSXXXXIEHVESVEPYLPYXXIDPOLFTKGASXLVKAALDTKQNKXEVQLR 180
DB 121 IVEYEGDVKSATSTIEHVESVEPYLPYRIDPOLFTKGASXLVKAALDTKQNKXEVQLR 180
QY 181 GIEKIAQXXXXNDVYITAKPEYKVMNDVARGIVADVAOSSYGLYGGQGIYAVALDTGLD 240
DB 181 GIEKIAQFAISNDVYITAKPEYKVMNDVARGIVADVAOSSYGLYGGQGIYAVALDTGLD 240
QY 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGHVAGSVLGNSTNGKGAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGHVAGSVLGNSTNGKGAPOANLVF 300
QY 301 QSIMDSXGGLGGLPSNLQTLFSGQASAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDM 360
DB 301 QSIMDSXGGLGGLPSNLQTLFSGQASAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDM 360

```

```

Db      301 OSIMDSGGIGLPSNLTQLFSQAYSAGARIHTNSWGAAMGAYTTRSDRVDDYAKNDM 360
QY      361 TLIFAGNEXPNPGTISAPGTAKNATVGAENTLRPSGSYADINIHVAQFSSRGPTXOG 420
Db      361 TLIFAGNEXPNPGTISAPGTAKNATVGAENTLRPSGSYADINIHVAQFSSRGPTXOG 420
QY      421 RIKPDVMAQGTILISARSSILAPDSSFWANHDSKYAYMGTSMAATPIVAGNVAQLREHFYK 480
Db      421 RIKPDVMAQGTILISARSSILAPDSSFWANHDSKYAYMGTSMAATPIVAGNVAQLREHFYK 480
QY      481 NNGITPKESLKAALJAGAADXGLGYPNGNQGWRVTLDKSLNVAAYNESSSLSTSOKAT 540
Db      481 NNGITPKESLKAALJAGAADXGLGYPNGNQGWRVTLDKSLNVAAYNESSSLSTSOKAT 540
QY      541 YKFTATAGKPLKISLWSDAPASTASTVTLVNDLDTVTPANGTXYVGDPEKXPXKMD 600
Db      541 YKFTATAGKPLKISLWSDAPASTASTVTLVNDLDTVTPANGTXYVGDPEKXPXKMD 600
QY      601 GRNNVENVFINKPQSGTYTIEVOAYNVVPGQXFSIAIYN 640
Db      601 GRNNVENVFINKPQSGTYTIEVOAYNVVPGQXFSIAIYN 640

RESULT 6
AA117089
ID      AA117089 standard; protein; 639 AA.
AC      AA117089;
XX
XX      20-MAR-2003 (revised)
DT      21-JUL-1999 (first entry)
XX
XX      Bacillus alkaline protease.
DE
KW      Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW      washing composition; oxidising agent.
XX
XX      Bacillus sp.
OS
XX      WO9918218-A1.
XX
XX      15-APR-1999.
XX
XX      07-OCT-1998; 98WO-JP004528.
XX
XX      07-OCT-1997; 97JP-00274570.
XX
XX      (KAOs ) KAO CORP.
XX
XX      Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI      Shikata S, Nomura M;
XX
XX      WPI, 1999-287736/27.
DR      N-PSDB; AAX37277.
XX
XX      Alkali protease from Bacillus used in washing powders.
XX
XX      Disclousure; Page 53-58; 71pp; Japanese.
XX
XX      The invention relates to alkaline proteases produced by strains of
CC      Bacillus. The proteases ability to digest casein is not inhibited by
CC      oleic acid and they have a high stability to oxidising agents. The
CC      alkaline protease of the invention has the following properties: (a) it
CC      is active over the pH range 4-13 and has at least 80% of its optimum
CC      activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC      stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC      its ability to digest casein is not inhibited by oleic acid; (e) it has
CC      molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC      used as enzymes in washing compositions for use in automatic dishwashers
CC      and for washing clothes. The stability to oxidising agents allows the
CC      enzyme to be an effective component of washing compositions including
CC      bleaches. The present sequence represents an alkaline protease. (Updated
CC      on 20-MAR-2003 to correct DR field.)

```

```

XX      XX
SQ      Sequence 639 AA;
Query Match      98.7%; Score 3020; DB 2; Length 639;
Best Local Similarity 93.4%; Pred. No. 8.7e-237;
Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY      4 KKKVFLSVTSAALISTVALXNPISAGAXXFLDLPFGIGTDTTDXGFSKXOTGAAPFY 63
Db      3 KKKVFLSVTSAALISTVALXNPISAGAXXFLDLPFGIGTDTTDXGFSKXOTGAAPFY 62
QY      64 ESENVTLKGLKKKLETPANNKLIHXFNGPILBETKQXLEXTGAKLIDYIPDYAYIV 123
Db      63 ESENVTLKGLKKKLETPANNKLIHXFNGPILBETKQXLEXTGAKLIDYIPDYAYIV 122
QY      124 YEGDYKSKXXXIEHVESVPEYLPYXXIDPQLFTKASXIVKAXALDTQXNKEVQLRGIE 183
Db      123 YEGDVQSKVRSIEHVESVPEYLPYKIDPQLFTKASITLVKALALDTQNNKEVQLRGIE 182
QY      184 XIAQXXXNDVYVITAKPEYKYMNDVARGIVKADVAQSSYGLYGQQLVAVADTGLDTR 243
Db      183 EIAQVYASNDVYVITAKPEYKYMNDVARGIVKADVAQSSYGLYGQQLVAVADTGLDTR 242
QY      244 NDSWHEAFRGKITLALYALGRTNNANDTNGSTHVAAGTLAGNXTKGMAPQANTVFGSI 303
Db      243 NDSWHEAFRGKITLALYALGRTNNANDTNGSTHVAAGTLAGNXTKGMAPQANTVFGSI 302
QY      304 MDSXGGLGGLPSNLTQLFSQAYSAGARIHTNSWGAAMGAYTTRSDRVDDYAKNDMTIL 363
Db      303 MDSXGGLGGLPSNLTQLFSQAYSAGARIHTNSWGAAMGAYTTRSDRVDDYAKNDMTIL 362
QY      364 PAAQNEXPNGGTISAPGTAKNATVGAENTLRPSGSYADINIHVAQFSSRGPTXOGRIK 423
Db      363 PAAQNEXPNGGTISAPGTAKNATVGAENTLRPSGSYADINIHVAQFSSRGPTXOGRIK 422
QY      424 PVMAAPGTIISARSSILAPDSSFWANHDSKYAYMGTSMAATPIVAGNVAQLREHFYKRG 483
Db      423 PVMAAPGTIISARSSILAPDSSFWANHDSKYAYMGTSMAATPIVAGNVAQLREHFYKRG 482
QY      484 ITPKPSLKAALJAGAADXGLGYPNGNQGWRVTLDKSLNVAAYNESSSLSTSOKATYYP 543
Db      483 ITPKPSLKAALJAGAADXGLGYPNGNQGWRVTLDKSLNVAAYNESSSLSTSOKATYYP 542
QY      544 TATAGKPLKISLWSDAPASTASTVTLVNDLDTVTPANGTXYVGDPEKXPXKMDGKN 603
Db      543 TATAGKPLKISLWSDAPASTASTVTLVNDLDTVTPANGTXYVGDPEKXPXKMDGKN 602
QY      604 NVENVFINKPQSGTYTIEVOAYNVVPGQXFSIAIYN 640
Db      603 NVENVFINKPQSGTYTIEVOAYNVVPGQXFSIAIYN 639

RESULT 7
AA117088
ID      AA117088 standard; protein; 640 AA.
AC      AA117088;
XX
XX      20-MAR-2003 (revised)
DT      21-JUL-1999 (first entry)
XX
XX      An alkaline protease sequence from Bacillus species.
DE
KW      Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW      washing composition; oxidising agent.
XX
XX      Bacillus sp.
OS
XX
XX      Key      Location/Qualifiers
FH      Misc-difference 1..640
FT      /note="all residues indicated as Xaa are arbitrary amino
FT      acids"
XX

```



```

PN WO9918218-A1.
XX 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX (KAOS ) KAO CORP.
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
XX Shikata S, Nomura M;
XX
XX WPI; 1999-287736/27.
XX N-PSDB; AAX37278.
XX
XX Alkali protease from Bacillus used in washing powders.
XX
XX Claim 3; Page 50-53; 71pp; Japanese.
XX
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease of the
CC invention. (Updated on 20-MAR-2003 to correct DR field.)
XX
XX
XX Sequence 640 AA;
SQ
Query Match          98.6%; Score 3016; DB 2; Length 640;
Best Local Similarity 100.0%; Pred. No. 1.8e-236;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRXKKKVFSLVLSAAAILSTVALXNPSAGXARXFDLDFKGIOTTTDXGFSKXQXTGAAA 60
DB 1 MRXKKKVFSLVLSAAAILSTVALXNPSAGXARXFDLDFKGIOTTTDXGFSKXQXTGAAA 60
QY 61 FLVSEENVTLKXGKLGKLETPANNKLTIXQFNGPILEETKQXLEXTAKIIDYIPDYAY 120
DB 61 FLVSEENVTLKXGKLGKLETPANNKLTIXQFNGPILEETKQXLEXTAKIIDYIPDYAY 120
QY 121 IVEYEGDVYKXXXXIEHVESVEPYLPXYXIDPQLFTKGASXIVKAXALDTKXNREVOLR 180
DB 121 IVEYEGDVYKXXXXIEHVESVEPYLPXYXIDPQLFTKGASXIVKAXALDTKXNREVOLR 180
QY 181 GIEIXAQXXXXSNDVXYITAKPEYKVMNDVARGIVRADVAQSSYGLYGQGOIYAVADTGLD 240
DB 181 GIEIXAQXXXXSNDVXYITAKPEYKVMNDVARGIVRADVAQSSYGLYGQGOIYAVADTGLD 240
QY 241 TGRNDSMHEAFRGITTLALYALGRTNNADTNGHTHVA GSTLGNKTKKGAPOANLVF 300
DB 241 TGRNDSMHEAFRGITTLALYALGRTNNADTNGHTHVA GSTLGNKTKKGAPOANLVF 300
QY 301 QSIMDSXGGLGGLPNTLQTLFQOAXSAGARIRHTNSMGAAVANGAYTTDSRNVDDYVRKNDM 360
DB 301 QSIMDSXGGLGGLPNTLQTLFQOAXSAGARIRHTNSMGAAVANGAYTTDSRNVDDYVRKNDM 360
QY 361 TILFAAGEXEPNGGITSAPGTAKNAITVGATENLPPSFGSVADNINHVAFSSRSGPTXDG 420
DB 361 TILFAAGEXEPNGGITSAPGTAKNAITVGATENLPPSFGSVADNINHVAFSSRSGPTXDG 420
QY 421 RIKPVMVAPGTIILARSLSLADSSFMANHDSKYVMGTSNATPIVAGNVAQLREHVPK 480
DB 421 RIKPVMVAPGTIILARSLSLADSSFMANHDSKYVMGTSNATPIVAGNVAQLREHVPK 480

```

```

QY 481 NRGITPKPSLILKALILAGAADXGLGYPNGNGMGRVTLDDKSLNAVYVNESSXLSTSOXAT 540
DB 481 NRGITPKPSLILKALILAGAADXGLGYPNGNGMGRVTLDDKSLNAVYVNESSXLSTSOXAT 540
QY 541 YXFTATAGKPKIKISLWSDAPASTTASVTLVNDLDTITAPNGTYTYGNDPFXEXXXKND 600
DB 541 YXFTATAGKPKIKISLWSDAPASTTASVTLVNDLDTITAPNGTYTYGNDPFXEXXXKND 600
QY 601 GRNNVENVFINKPQSGTITTEVOAYNVYPVGQXFSLAIYN 640
DB 601 GRNNVENVFINKPQSGTITTEVOAYNVYPVGQXFSLAIYN 640

RESULT 8
AA17087
ID AA17087 standard; protein; 639 AA.
XX
XX AA17087;
XX
XX 20-MAR-2003 (revised)
XX 21-JUL-1999 (first entry)
XX
XX An alkaline protease sequence from Bacillus species.
XX
XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX washing composition; oxidizing agent.
XX
XX Bacillus sp.
XX
XX Key Location/Qualifiers
FH Misc-difference 1..639
FT /note= "all residues indicated as Xaa are arbitrary amino
FT acids"
XX
XX
XX WO9918218-A1.
XX
XX 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX (KAOS ) KAO CORP.
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
XX Shikata S, Nomura M;
XX
XX WPI; 1999-287736/27.
XX N-PSDB; AAX37277.
XX
XX Alkali protease from Bacillus used in washing powders.
XX
XX Claim 3; Page 47-50; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
XX Bacillus. The proteases ability to digest casein is not inhibited by
XX oleic acid and they have a high stability to oxidizing agents. The
XX alkaline protease of the invention has the following properties: (a) it
XX is active over the pH range 4-13 and has at least 80% of its optimum
XX activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
XX stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
XX its ability to digest casein is not inhibited by oleic acid; (e) it has
XX molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
XX used as enzymes in washing compositions for use in automatic dishwashers
XX and for washing clothes. The stability to oxidizing agents allows the
XX enzyme to be an effective component of washing compositions including
XX bleaches. The present sequence represents an alkaline protease of the
XX invention. (Updated on 20-MAR-2003 to correct DR field.)
XX
XX
XX Sequence 639 AA;
SQ
Query Match          98.3%; Score 3007; DB 2; Length 639;
Best Local Similarity 100.0%; Pred. No. 9.9e-236;

```

Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 4 KKKVPLSVLSAAAIISTVALXNPSAGAXRFDLDFKGIQTTTDXKFSKQXQGAAPLV 63
DB 3 KKKVPLSVLSAAAIISTVALXNPSAGAXRFDLDFKGIQTTTDXKFSKQXQGAAPLV 62
QY 64 ESENVLXKGLKKEETVPANNKLAHXQFNGLFEETKOKLXETGAKILDYIPDYAYIE 123
DB 63 ESENVLXKGLKKEETVPANNKLAHXQFNGLFEETKOKLXETGAKILDYIPDYAYIE 122
QY 124 YEGDVYKXXXXXIEHVESVEPYLPXYXIDPQLFTKASXLYKAXALDTKQXNKEVQURGIE 183
DB 123 YEGDVYKXXXXXIEHVESVEPYLPXYXIDPQLFTKASXLYKAXALDTKQXNKEVQURGIE 182
QY 184 XIAQXXXXNDVYITTAPEYKVMNDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTR 243
DB 183 XIAQXXXXNDVYITTAPEYKVMNDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTR 242
QY 244 NDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLGNKXTNKGMAPOANLVFQSI 303
DB 243 NDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLGNKXTNKGMAPOANLVFQSI 302
QY 304 MDSXGGLGGLPSNLQTLFQAXSAGARITHNSWGAAVNGAYTTDSRVVDYVRKNDMTIL 363
DB 303 MDSXGGLGGLPSNLQTLFQAXSAGARITHNSWGAAVNGAYTTDSRVVDYVRKNDMTIL 362
QY 364 PAAGNEXPNVGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAFQSSRGPTKGRIK 423
DB 363 PAAGNEXPNVGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAFQSSRGPTKGRIK 422
QY 424 PVVMAPGTITLSARSSILAPDSSFWANHDSKAYMGSTSMATPIVGNVVAQLREHPKNG 483
DB 423 PVVMAPGTITLSARSSILAPDSSFWANHDSKAYMGSTSMATPIVGNVVAQLREHPKNG 482
QY 484 ITPKPSLLKALIAAGADKGLGYPNGNQMGRTVLDKSLNVAVYNESSXILSTOKATYXF 543
DB 483 ITPKPSLLKALIAAGADKGLGYPNGNQMGRTVLDKSLNVAVYNESSXILSTOKATYXF 542
QY 544 TATACKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXYVGNDFYKPKXXNWDGN 603
DB 543 TATACKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXYVGNDFYKPKXXNWDGN 602
QY 604 NVENVFINKPOSGCTTIEVQAVNVPGQXPSLAVN 640
DB 603 NVENVFINKPOSGCTTIEVQAVNVPGQXPSLAVN 639

```

RESULT 9

AAW89547 standard; protein; 641 AA.

```

AC AAW89547;
XX
XX 12-APR-1999 (first entry)
XX
DB Bacillus JPI170 protease.
XX
KM Protease; detergent; surfactant; leather processing; debittering;
XX flavour.
XX
XX Bacillus sp.
XX
XX Key 1. Location/Qualifiers
XX Peptide 1. .33
XX FT /note="signal peptide"
XX FT 34. .208
XX FT /note="prepro region"
XX FT 209. .641
XX FT /note="mature protein"
XX
XX MO9856927-A2.
XX
XX 17-DEC-1998.

```

XX 09-JUN-1998; 98WO-US012005.
 PF 12-JUN-1997; 97US-00873479.
 PR (NOVO) NOVO NORDISK BIOTECH INC.
 PA Sloma A. Christensen L;
 PI WPI; 1999-080908/07.
 DR N-PSDB; AAW82382.
 XX
 PT Novel protease from *Bacillus subtilis* LC20 - useful in laundry and
 dishwashing detergents and for leather processing.
 Claim 7; Page 53-54; 77pp; English.

This is the amino acid sequence of a novel protease of *Bacillus* sp. JPI170 (NCIB 12513), as deduced from the nucleotide sequence of an isolated gene (see AAW82382). The entire protein, including the signal peptide and prepro region, has 77% identity to alkaline protease Y (see AAW89548) from *Bacillus*. The invention provides vectors, recombinant host cells and methods for the recombinant production of the protease. The protease is used in laundry and dishwashing detergents, for institutional and industrial cleaning, and for leather processing, as well as for debittering and enhancing the degree of hydrolysis of protein hydrolyzates, for flavour development through hydrolysis of peptides, degradation of undesired peptides and in enzymatic synthesis of peptides. It has enhanced stability towards oxidation under alkaline conditions, e.g. towards bleaching agents of the peroxy type. The invention also provides mutant cells in which the protease activity is diminished. Such cells can be used for the production of heterologous recombinant proteins

SQ Sequence 641 AA;

Query Match 89.0%; Score 2722.5; DB 2; Length 641;

Best Local Similarity 82.9%; Pred. No. 1.4e-212; Matches 532; Conservative 40; Mismatches 67; Indels 3; Gaps 2;

```

QY 1 MKX--KVVFLSVLSAAAIISTVALXNPSAGAXRFDLDFKGIQTTTDXKFSKQXQGA 58
DB 1 MKKSKVFLSVLSVAALLSSVALSSPBTGANNFELDPKIETLTLEKATKQKTK 60
QY 59 AAFVSENVKXKGLKKEETVPANNKLAHXQFNGLFEETKOKLXETGAKILDYIPDY 118
DB 61 ASFLVSENVKPKSIQKKEVVPADNKLXYIQFGPILEFQTLQLEKTKGAKILDYIPDY 120
QY 119 AYIYVEEDVYKXXXXXIEHVESVEPYLPXYXIDPQLFTKASXLYKAXALDTKQXNKEVQ 178
DB 121 AYIYVEDDVAQVNTAIAHLSVEVEPYLPXYXIDPQLFSGASELVEYVALDKKORSKEVR 180
QY 179 LRGIEXIAQXXXXNDVYITTAPEYKVMNDVARGIVKADVAQSSYGLYGQGIYAVADT 238
DB 181 LRGLAQIAQYATNDVLYITTAPEYKVMNDVARGIVKADVAQNNGLYGQGIYAVADT 240
QY 239 LDTGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLGNKXTNKGMAPOANT 298
DB 241 LDTGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLGNKXTNKGMAPOANT 299
QY 299 VFGSIVDSXGGLGGLPSNLQTLFQAXSAGARITHNSWGAAVNGAYTTDSRVVDYVRKN 358
DB 300 VFGSIVDSXGGLGGLPSNLQTLFQAXSAGARITHNSWGAAVNGAYTTDSRVVDYVRKN 359
QY 359 DMTILFAAGNEXPNVGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAFQSSRGPTK 418
DB 360 DMTILFAAGNEXPNVGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAFQSSRGPTK 419
QY 419 DGRIPDVMAPGTITLSARSSILAPDSSFWANHDSKAYMGSTSMATPIVAGNVQALREHF 478
DB 420 DGRIPDVMAPGTITLSARSSILAPDSSFWANHDSKAYMGSTSMATPIVAGNVQALREHF 479
QY 479 VKNRGITTPKPSLLKALIAAGADKGLGYPNGNQMGRTVLDKSLNVAVYNESSXILSTOK 538

```

Db 480 VKRGVTPKPSILKALALAGADVGLGFPNGNGRGVTLDKSLANVAFNETSPLSTSQK 539
Qy 539 ATTYFTATAGKPLKISLVMSDAPASTTASVTLVNDLDVITAPNGTXYVGNDFXXPKXXN 598
Db 540 ATYSPTAAGKPLKISLVMSDAPGSTTASLTIVNDLDVITAPNGTKYVGNDFTPAYDNN 599
Qy 599 WGRNNVENVFINKPOSGTYTIEVOAYNVVPGPOKFSALIVN 640
Db 600 WGRNNVENVFINKPOSGTYTIEVOAYNVVPGPOKFSALIVH 641

RESULT 10
AAM89548 standard; protein; 636 AA.
AAM89548;
12-APR-1999 (first entry)
Bacillus sp. alkaline protease Y.
Alkaline protease Y; detergent; surfactant; leather processing; deblotting; flavour.
Bacillus sp.
MO9856927-A2.
17-DEC-1998.
09-JUN-1998; 98WO-US012005.
12-JUN-1997; 97US-00873479.
(NOVO) NOVO NORDISK BIOTECH INC.
Sloma A, Christensen L;
WPI; 1999-080908/07.
Novel protease from Bacillus subtilis LC20 - useful in laundry and dishwashing detergents and for leather processing.
Claim 3; Page 55-56; 77pp; English.
This is the amino acid sequence of a Bacillus sp. alkaline protease Y that is said to have good alkali and surfactant resistance and improved degergency. It shows 77% identity to a newly isolated protease (see AAM89547) of Bacillus sp. JP170 (NCIB 12513). The invention provides vectors, recombinant host cells and methods for the recombinant production of such proteases. The protease are used in laundry and dishwashing detergents, for institutional and industrial cleaning, and for leather processing, as well as for deblotting and enhancing the degree of hydrolysis of protein hydrolysates, for flavour development through hydrolysis of proteins, degradation of undesired peptides and in enzymatic synthesis of peptides. They have enhanced stability towards oxidation under alkaline conditions, e.g. towards bleaching agents of the peroxy type. The invention also provides mutant cells in which the protease activity is diminished. Such cells can be used for the production of heterologous recombinant proteins

Query Match 78.9%; Score 2414; DB 2; Length 636;
Best Local Similarity 73.0%; Pred. No. 1.8e-187;
Matches 468; Conservative 68; Mismatches 99; Indels 6; Gaps 5;

Qy 1 MRKKKVVLSVSAALISTVALKPSAGXAXFPDLDFKIGITTTDDXXGFSQXOTGAAA 60
Db 1 MKGKKRVVLSVSAALISTVALKPSAGXAXFPDLDFKIGITTTDDXXGFSQXOTGAAA 57
Qy 61 FLVSENVKTKKGLKKKLETVANNKMLHXOFN-GPILEETKQXLEXTGAKILDIYIPDYA 119

Db 58 FLVDNTENINIPKGIQKTLKLVNQKONELYIQFTGPTPISEERKGLSEGLSDIVDPDYA 117
Qy 120 YIYIEGVDKSKXXXIEHVESVEPYLPXYIDPOLFTKGASXLVYKAAALDTKQXKEVOL 179
Db 118 FIVQYSG-ATKNISITLSHSEVENVQFPLPYKIDPELITKGASQLVQAVLANKHEKKNKF 176
Qy 180 RGIEKIXQXXXSNVXYITAKEEYKMNNDVARGIVKADVAAQSSGLVGGQGIIVAVADTGL 239
Db 177 TGLSEIVQYAAANDVLYISPKPEYELMNDVARGIVKADVAAQSNYGLVGGQGLVAVADTGL 236
Qy 240 DTGRNDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVLGNXTNKGMAPOANLV 299
Db 237 DTGRNDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVLGN-ALANKGMAPOANLV 295
Qy 300 FQSTMDXKGLIGLPSNLQTLFQDAXSAGARIHTNSGAAVNGAYTTDSRNVDDYRKND 359
Db 296 FQSTMDXKGLIGLPSNLQTLFQDAXSAGARIHTNSGAAVNGAYTTDSRNVDDYRKND 355
Qy 360 MTLLPAAGNEKPNNGTISAPGTAKNAITVGATENTLPPSGSYADININVAQFSSRGPTKD 419
Db 356 MTVLPAAGNEKPNNGTISAPGTAKNAITVGATENTLPPSGSYADININVAQFSSRGATRD 415
Qy 420 GRIPDYVAPGTXILSARSSLPDSSFMANHDSKYAVWGTSMATPIVAGNVADLREHFV 479
Db 416 GRIPDYVAPGTXILSARSSLPDSSFMANHDSKYAVWGTSMATPIVAGNVADLREHFV 475
Qy 480 KNRGITPKPSILKALALAGADVGLGFPNGNGRGVTLDKSLANVAYNBSXKSTSGKA 539
Db 476 KNRGITPKPSILKALALAGADVGLGFPNGNGRGVTLDKSLANVAYNBSXKSTSGKA 535
Qy 540 TYXFPATGKPLKISLVMSDAPASTTASVTLVNDLDVITAPNGTXYVGNDFXXPKXXNW 599
Db 536 TYXFPATGKPLKISLVMSDAPASTTASVTLVNDLDVITAPNGTXYVGNDFXXPKXXNW 595
Qy 600 DGRNNVENVFINKPOSGTYTIEVOAYNVVPGPOKFSALIVN 640
Db 596 DGRNNVENVFINKPOSGTYTIEVOAYNVVPGPOKFSALIVH 636

RESULT 11
AAM50080 standard; protein; 434 AA.
AAM50080;
12-AUG-2002 (first entry)
Bacillus sp KSM-KP43 alkaline protease protein fragment.
Alkaline protease; detergent; laundry; bleaching; dishwasher.
Bacillus sp.
EP1209233-A2.
29-MAY-2002.
22-NOV-2001; 2001EP-00127851.
22-NOV-2001; 2000JP-00355166.
12-APR-2001; 2001JP-00114048.
(KAOS) KAO CORP.
Hacada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N,
Pi Okuda M, Saeki K;
WPI; 2002-437516/47.
New modified alkaline proteases useful in detergent compositions.
Claim 1; Page 10-11; 25pp; English.

CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency %
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease KP43 from
 CC *Bacillus* sp strain KSM-KP43 which is used to create the modified protease
 CC represented in AAM50081

XX
 XX
 SQ Sequence 434 AA;

Query Match 70.4%; Score 2155; DB 5; Length 434;
 Best Local Similarity 96.3%; Pred. No. 1.3e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 207 NDVAGIVKADVAQSSYGLYGGQIVAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 266
 DB 1 NDVAGIVKADVAQSSYGLYGGQIVAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 60
 QY 267 NANDTNGHGTTHVAGSVLGNKTKMGMAPOANLVFOSIMDSXGGLGSLPSNLQTLFSSQAXS 326
 DB 61 NANDTNGHGTTHVAGSVLGNKTKMGMAPOANLVFOSIMDSXGGLGSLPSNLQTLFSSQAXS 120
 QY 327 AGARIHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILPAAGNEXPNGGTISAPGTAKNAI 386
 DB 121 AGARIHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILPAAGNEXPNGGTISAPGTAKNAI 180
 QY 387 TVGATENLRPSFGSYADININHAQSSSRGPTDGRIKPDVMAPGTYIISARSSLPDSSSF 446
 DB 181 TVGATENLRPSFGSYADININHAQSSSRGPTDGRIKPDVMAPGTYIISARSSLPDSSSF 240
 QY 447 WANHDSKYAYMGSTMATPIVAGNVAQLREHFVKNRGITPPKPSLKALIALGADVXGLCY 506
 DB 241 WANHDSKYAYMGSTMATPIVAGNVAQLREHFVKNRGITPPKPSLKALIALGADVXGLCY 300
 QY 507 PNGNGMGRVTLDKSLNVAAYNESXLSSTOKATYXFTATAGKPKLISLWSDAPASTTA 566
 DB 301 PNGNGMGRVTLDKSLNVAAYNESXLSSTOKATYXFTATAGKPKLISLWSDAPASTTA 360
 QY 567 SVTLVNDLDLVITAPNGRTYVGNDFXKPKXKXWDRNNVENVFINKPQSGTYTIEVOAYN 626
 DB 361 SVTLVNDLDLVITAPNGRTYVGNDFXKPKXKXWDRNNVENVFINKPQSGTYTIEVOAYN 420
 QY 627 VPVGPOKFSLAIVN 640
 DB 421 VPVGPOKFSLAIVN 434

RESULT 12
 AAM50081

ID AAM50081 standard; protein; 434 AA.

XX
 AC AAM50081;

XX 12-AUG-2002 (first entry)

XX DE *Bacillus* sp KSM-KP9860 alkaline protease protein fragment.

XX KW Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX OS *Bacillus* sp.

XX PN EP1209233-A2.

XX PD 29-MAY-2002.

XX PF 22-NOV-2001; 2001EP-00127851.

XX PR 22-NOV-2000; 2000JP-00355156.

XX PR 12-APR-2001; 2001JP-00114048.

XX PA (KAO) KAO CORP.

XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

PI Okuda M, Saeki K;
 XX
 DR WPI; 2002-437518/47.

XX PT New modified alkaline proteases useful in detergent compositions.

XX PS Claim 5; Page 12-13; 25pp; English.

XX
 CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency %
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease KP9860 from
 CC *Bacillus* sp strain KSM-KP9860 described in the method of the invention

XX
 SQ Sequence 434 AA;
 Query Match 70.4%; Score 2155; DB 5; Length 434;
 Best Local Similarity 96.3%; Pred. No. 1.3e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 207 NDVAGIVKADVAQSSYGLYGGQIVAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 266
 DB 1 NDVAGIVKADVAQSSYGLYGGQIVAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 60
 QY 267 NANDTNGHGTTHVAGSVLGNKTKMGMAPOANLVFOSIMDSXGGLGSLPSNLQTLFSSQAXS 326
 DB 61 NANDTNGHGTTHVAGSVLGNKTKMGMAPOANLVFOSIMDSXGGLGSLPSNLQTLFSSQAXS 120
 QY 327 AGARIHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILPAAGNEXPNGGTISAPGTAKNAI 386
 DB 121 AGARIHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILPAAGNEXPNGGTISAPGTAKNAI 180
 QY 387 TVGATENLRPSFGSYADININHAQSSSRGPTDGRIKPDVMAPGTYIISARSSLPDSSSF 446
 DB 181 TVGATENLRPSFGSYADININHAQSSSRGPTDGRIKPDVMAPGTYIISARSSLPDSSSF 240
 QY 447 WANHDSKYAYMGSTMATPIVAGNVAQLREHFVKNRGITPPKPSLKALIALGADVXGLCY 506
 DB 241 WANHDSKYAYMGSTMATPIVAGNVAQLREHFVKNRGITPPKPSLKALIALGADVXGLCY 300
 QY 507 PNGNGMGRVTLDKSLNVAAYNESXLSSTOKATYXFTATAGKPKLISLWSDAPASTTA 566
 DB 301 PNGNGMGRVTLDKSLNVAAYNESXLSSTOKATYXFTATAGKPKLISLWSDAPASTTA 360
 QY 567 SVTLVNDLDLVITAPNGRTYVGNDFXKPKXKXWDRNNVENVFINKPQSGTYTIEVOAYN 626
 DB 361 SVTLVNDLDLVITAPNGRTYVGNDFXKPKXKXWDRNNVENVFINKPQSGTYTIEVOAYN 420
 QY 627 VPVGPOKFSLAIVN 640
 DB 421 VPVGPOKFSLAIVN 434

RESULT 13

ID ADY33778 standard; protein; 434 AA.

XX
 AC ADY33778;

XX DT 05-MAY-2005 (first entry)

XX DE *Bacillus* species alkaline protease.

XX KW mutagenesis; alkaline protease; detergent; laundry; cosmetic; food;
 XX bleaching detergent.

XX OS *Bacillus* sp.

XX PN EP1347044-A2.

XX PD 24-SEP-2003.

XX

21-MAR-2003; 2003EP-00006472.
 22-MAR-2002; 2002JP-00081428.
 06-JUN-2002; 2002JP-00165987.
 18-OCT-2002; 2002JP-00304230.
 18-OCT-2002; 2002JP-00304231.
 (KAO) KAO CORP.
 Okuda M, Sato T, Saito K, Sumitomo N, Izawa Y, Saeki K;
 Kobayashi T, Nomura M;
 WPI; 2003-846540/79.
 N-PSDB; ADY33779.
 New alkaline protease having specific amino acid residue at a specific position of its amino acid sequence, useful for producing detergent compositions, laundry detergent, fiber modifiers, leather-treating agents or pipe cleaners.
 Claim 1; SEQ ID NO 1; 31pp; English.
 The invention relates to an alkaline protease having a fully defined sequence of 434 amino acids (I) given in the specification, or an amino acid sequence at least 80% homology with (I), where an amino acid residue at position 65, 101, 163, 170, 171, 273, 320, 359 or 387 of (I) is selected from 46 amino acid residues. The amino acid residues at the corresponding positions are selected from: position 65: proline; position 101: asparagine; position 163: histidine, aspartic acid, phenylalanine, lysine, asparagine, serine, isoleucine, leucine, glutamine, threonine or valine; position 170: valine or leucine; position 171: alanine, glutamic acid, glycine or threonine; position 273: isoleucine, glycine or threonine; position 320: phenylalanine, valine, threonine, leucine, isoleucine or glycine; position 359: serine, leucine, valine, isoleucine or glutamic acid; and position 387: alanine, lysine, glutamine, glutamic acid, arginine or histidine. The alkaline protease is useful for the production of a detergent composition, such as laundry detergent, fiber modifiers, leather-treating agents, cosmetic compositions, bath additives, food modifiers and pharmaceutical compositions. The alkaline protease may also be used as bleaching detergent, hard surface cleansing detergent, pipe cleaner, artificial tooth cleaner, and as a sterilizing cleanser for medical tools. The new alkaline protease has a more potent proteolytic capacity, exhibiting excellent detergency for the removal of a complex stain, and has high secretion capacity. This sequence corresponds to the Bacillus sp. alkaline protease.
 Sequence 434 AA;
 Query Match 70.4%; Score 2155; DB 7; Length 434;
 Best Local Similarity 96.3%; Pred. No. 1.3e-166; Indels 0; Gaps 0;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 207 NDVARGIVKADVAOSSYGLYGGQIVAAVDGLDGRNDSSMHEAFRKITLALYALGRTN 266
 1 NDVARGIVKADVAOSSYGLYGGQIVAAVDGLDGRNDSSMHEAFRKITLALYALGRTN 60
 267 NANTNGHGTTHVAGSVLNGKGTNKGMAPOANVFGOSINDSGKGLGSPNLQTLPSQAXS 326
 61 NANTNGHGTTHVAGSVLNGKGTNKGMAPOANVFGOSINDSGKGLGSPNLQTLPSQAXS 120
 327 AGARLHTMSWGAANVGAAYTTDSRNVDDYRKADMTLLFAAGNEXPNGTISAPGAKNAI 386
 121 AGARLHTMSWGAANVGAAYTTDSRNVDDYRKADMTLLFAAGNEXPNGTISAPGAKNAI 180
 387 TVGATRNLRPSFGSYADNINHAQPSRSRPTKGRIKPDVMAPTXILSARSLAPDSGF 446
 121 TVGATRNLRPSFGSYADNINHAQPSRSRPTKGRIKPDVMAPTXILSARSLAPDSGF 240
 447 WANHDSKTAVMGTSMTPTIVAGNVAQLREHFVYKRGITPKSLKALAIAGAAXIGCY 506
 241 WANHDSKTAVMGTSMTPTIVAGNVAQLREHFVYKRGITPKSLKALAIAGAAXIGCY 300
 507 PNGNGMGKRVTLDSKLVNVAVNESSSLTSGKATYXFTAPAKPLKISLVMSDAPASTTA 566

301 PNGNGMGKRVTLDSKLVNVAVNESSSLTSGKATYXFTATNGKPLKISLVMSDAPASTTA 360
 567 SVTLVNDLDTVITAPNGTXVYVNDPXXPKXKNDGRRNVNVPINXPGSGTYTIEVQAYN 626
 361 SVTLVNDLDTVITAPNGTXVYVNDPXXPKXKNDGRRNVNVPINXPGSGTYTIEVQAYN 420
 627 VPVGPQKFSALIVN 640
 421 VPVGPQKFSALIVN 434
 RESULT 14
 ADZ51758
 ID ADZ51758 standard; protein; 434 AA.
 ADZ51758;
 16-JUN-2005 (first entry)
 DE Mutant Bacillus sp. alkali protease, SEQ ID No:2.
 DE Mutant Bacillus sp. alkali protease, mutagenesis; mutein.
 KM alkali protease; enzyme; surfactant; mutagenesis; mutein.
 XX Bacillus sp.
 OS Synthetic.
 XX JP2003125783-A.
 PD 07-MAY-2003.
 XX 26-OCT-2001; 2001JP-00329472.
 PF 26-OCT-2001; 2001JP-00329472.
 XX 26-OCT-2001; 2001JP-00329472.
 FR (KAO) KAO CORP.
 PA
 XX WPI; 2003-856569/80.
 DR
 PT New alkali protease useful for preparing detergents comprises
 PT substitution mutations at 251 or 256 position of protease KP43 derived
 from Bacillus species KSM-KP43.
 XX
 PS Claim 3; SEQ ID NO 2; 16pp; Japanese.
 XX
 CC The invention relates to a mutant alkali protease having an amino acid
 CC deletion at position(s) 251 or 256 in a fully defined sequence given as
 CC SEQ ID No:1 in the specification, or the following amino acid residue
 CC Asp, Gly, Ala, Asp, Thr, Ile, Val, Leu or Glu at position 251, or the
 CC amino acid residue Lys, Ser, Phe, Arg, Tyr, Met, Thr, Ile, Val, Leu, or
 CC Glu at position 256. Also described are: (i) an alkali protease having to
 CC the amino acid sequence of SEQ ID No:1, or having 60% or more homology to
 CC SEQ ID No:1, or a deletion of amino acid corresponding to position 251 or
 CC 256, (ii) a gene which encodes an alkali protease, (iii) a recombinant
 CC vector containing the gene, and (iv) a transformed organism containing
 CC the recombinant vector. The alkali protease sequence having 60% or more
 CC homology to SEQ ID No:1 is chosen from a fully defined sequence selected
 CC from SEQ ID Nos 2-7 as given in the specification. The transformed
 CC organism is a microorganism. The mutant alkali protease is useful for
 CC preparing detergents. The mutant alkali protease exhibits high resistance
 CC against oxidizing agent. The mutant alkali protease has high specific
 CC activity. This sequence represents a mutant Bacillus sp. alkali protease.
 XX
 SQ Sequence 434 AA;
 Query Match 70.4%; Score 2155; DB 7; Length 434;
 Best Local Similarity 96.3%; Pred. No. 1.3e-166; Indels 0; Gaps 0;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 207 NDVARGIVKADVAOSSYGLYGGQIVAAVDGLDGRNDSSMHEAFRKITLALYALGRTN 266
 1 NDVARGIVKADVAOSSYGLYGGQIVAAVDGLDGRNDSSMHEAFRKITLALYALGRTN 60

QY 267 NANDTNGHGTTHVAGSVLGNKGTNKGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 326
 DB 61 NANDTNGHGTTHVAGSVLGNKGTNKGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 120
 QY 327 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEXPNGGTTISAPGTAKNAI 386
 DB 121 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEXPNGGTTISAPGTAKNAI 180
 QY 387 TVGATENTLRPSFGSYADNINHVAQPSRPTKDGRIKPDVMAFGTYILSARSSLAPDSSF 446
 DB 181 TVGATENTLRPSFGSYADNINHVAQPSRPTKDGRIKPDVMAFGTYILSARSSLAPDSSF 240
 QY 447 WANHSKTYAMGSGTSMATPIVAGNVAQLREHVVKNRGITPKPSLKAALIAAADIIGLGY 506
 DB 241 WANHSKTYAMGSGTSMATPIVAGNVAQLREHVVKNRGITPKPSLKAALIAAADIIGLGY 300
 QY 507 PNGNGMGRVTLDKSLNVAAYNNESSXLSTSOATYXFTATAGKPKLISLWSDAPASTTA 566
 DB 301 PNGNGMGRVTLDKSLNVAAYNNESSXLSTSOATYXFTATAGKPKLISLWSDAPASTTA 360
 QY 567 SYTLVNDLDLVTTAPNGTYVGNDFPKXPKXKXNDGRNNVENVFINKPQSGTTTIEVOAYN 626
 DB 361 SYTLVNDLDLVTTAPNGTYVGNDFPKXPKXKXNDGRNNVENVFINKPQSGTTTIEVOAYN 420
 QY 627 VPVGPOXPSLAIVN 640
 DB 421 VPVGPOXPSLAIVN 434
 RESULT 15
 ADZ51757
 ID ADZ51757 standard; protein; 434 AA.
 AC ADZ51757;
 DT 16-JUN-2005 (first entry)
 XX
 DE Wild-type *Bacillus* sp. alkali protease.
 XX alkali protease; enzyme; surfactant; mutagenesis.
 KM
 XX *Bacillus* sp.
 OS JP2003125783-A.
 PN 07-MAY-2003.
 PD 26-OCT-2001; 2001JP-00329472.
 XX 26-OCT-2001; 2001JP-00329472.
 PR 26-OCT-2001; 2001JP-00329472.
 XX (KAOS) KAO CORP.
 PA WPI; 2003-856569/80.
 DR
 XX
 PT New alkali protease useful for preparing detergents comprises
 PT substitution mutations at 251 or 256 position of protease KP43 derived
 PT from *Bacillus* species KSM-KP43.
 PS Claim 1; SEQ ID NO 1; 16pp; Japanese.
 XX
 CC The invention relates to a mutant alkali protease having an amino acid
 CC deletion at position(s) 251 or 256 in a fully defined sequence given as
 CC SEQ ID No:1 in the specification, or the following amino acid residue
 CC Asp, Gly, Ala, Asp, Thr, Ile, Val, Leu or Glu at position 251, or the
 CC amino acid residue Lys, Ser, Phe, Arg, Tyr, Met, Thr, Ile, Val, Leu, or
 CC Glu at position 256. Also described are: (i) an alkali protease having
 CC the amino acid sequence of SEQ ID No:1, or having 60% or more homology to
 CC SEQ ID No:1, or a deletion of amino acid corresponding to position 251 or
 CC 256, (ii) a gene which encodes an alkali protease, (iii) a recombinant
 CC vector containing the gene, and (iv) a transformed organism containing
 CC the recombinant vector. The alkali protease sequence having 60% or more
 CC homology to SEQ ID No:1 is chosen from a fully defined sequence selected

CC from SEQ ID Nos 2-7 as given in the specification. The transformed
 CC organism is a microorganism. The mutant alkali protease is useful for
 CC preparing detergents. The mutant alkali protease exhibits high resistance
 CC against oxidizing agent. The mutant alkali protease has high specific
 CC activity. This sequence represents wild-type *Bacillus* sp. alkali
 CC protease.
 XX
 SQ Sequence 434 AA;
 Query Match 70.4%; Score 2155; DB 7; Length 434;
 Best Local Similarity 96.3%; Pred. No. 1.3e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 207 NDVARGIVADVAQSSYGLYGQGIIVAAVDTGLDGRNDSWHEAFRGKITLVALGRTN 266
 DB 1 NDVARGIVADVAQSSYGLYGQGIIVAAVDTGLDGRNDSWHEAFRGKITLVALGRTN 60
 QY 267 NANDTNGHGTTHVAGSVLGNKGTNKGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 326
 DB 61 NANDTNGHGTTHVAGSVLGNKGTNKGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 120
 QY 327 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEXPNGGTTISAPGTAKNAI 386
 DB 121 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEXPNGGTTISAPGTAKNAI 180
 QY 387 TVGATENTLRPSFGSYADNINHVAQPSRPTKDGRIKPDVMAFGTYILSARSSLAPDSSF 446
 DB 181 TVGATENTLRPSFGSYADNINHVAQPSRPTKDGRIKPDVMAFGTYILSARSSLAPDSSF 240
 QY 447 WANHSKTYAMGSGTSMATPIVAGNVAQLREHVVKNRGITPKPSLKAALIAAADIIGLGY 506
 DB 241 WANHSKTYAMGSGTSMATPIVAGNVAQLREHVVKNRGITPKPSLKAALIAAADIIGLGY 300
 QY 507 PNGNGMGRVTLDKSLNVAAYNNESSXLSTSOATYXFTATAGKPKLISLWSDAPASTTA 566
 DB 301 PNGNGMGRVTLDKSLNVAAYNNESSXLSTSOATYXFTATAGKPKLISLWSDAPASTTA 360
 QY 567 SYTLVNDLDLVTTAPNGTYVGNDFPKXPKXKXNDGRNNVENVFINKPQSGTTTIEVOAYN 626
 DB 361 SYTLVNDLDLVTTAPNGTYVGNDFPKXPKXKXNDGRNNVENVFINKPQSGTTTIEVOAYN 420
 QY 627 VPVGPOXPSLAIVN 640
 DB 421 VPVGPOXPSLAIVN 434

Search completed: April 7, 2006, 17:37:31
 Job time : 139.608 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 17:37:51 ; Search time 27.5215 Seconds
(without alignments)
2237.477 Million cell updates/sec

Title: US-10-784-870-2
Perfect score: 3059
Sequence: 1 MRXKKKFLSVLSAAALST.....RYQAVNPVGPQXPSLAIYN 640

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Libbing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518.5	15.9	1743	2 T18279	multidrug resistan
2	477.5	15.6	1905	2 T18267	multidrug resistan
3	350.5	11.5	444	2 B83891	intracellular alka
4	319.5	10.4	442	2 A69587	intracellular alka
5	297.5	9.7	1398	2 T28159	pyrolysins (EC 3.4.
6	286	9.3	806	2 A41341	microbial serine p
7	279.5	9.1	419	1 S25835	subtilisin (EC 3.4
8	276.5	9.0	799	2 G83753	subtilisin-type pr
9	274.5	9.0	420	1 S23407	subtilisin (EC 3.4
10	270	8.8	580	2 S11890	serine proteinase
11	268.5	8.8	715	2 J04908	alkaline serine pr
12	262	8.6	894	2 P63730	cell wall-associat
13	260.5	8.5	1345	2 T29090	surface layer-asso
14	257.5	8.4	378	2 A33973	high-alkaline seri
15	255	8.3	513	1 A35742	aqualysin (EC 3.4.
16	250.5	8.2	380	2 A49778	serine proteinase
17	250.5	8.2	601	2 J04576	serine proteinase
18	250	8.2	382	2 J13780	subtilisin (EC 3.4
19	249.5	8.2	757	2 C84120	subtilisin-type pr
20	248	8.1	627	2 D75393	serine proteinase,
21	246.5	8.1	402	1 J00332	alkaline proteinas
22	245.5	8.0	381	2 J00778	subtilisin (EC 3.4
23	245	8.0	382	1 SUBSN	subtilisin (EC 3.4
24	243.5	8.0	381	1 SUBSI	subtilisin (EC 3.4
25	242.5	7.9	379	1 SUBSCL	subtilisin (EC 3.4
26	242	7.9	1331	2 A72647	probable surface 1
27	241.5	7.9	381	1 SUBSS	subtilisin (EC 3.4
28	239.5	7.8	401	2 J13974	serine proteinase
29	239.5	7.8	534	1 J06173	alkaline proteinase

30	237.5	7.8	381	2 J01487	subtilisin (EC 3.4
31	235	7.7	519	2 S71451	haloalysin R4 (EC 3
32	232	7.6	488	2 A11930	proteinnase (import
33	231.5	7.6	1167	1 A35066	streptococcal Csa
34	230	7.5	384	2 J04802	alkaline proteinase
35	223.5	7.3	409	1 S32905	serine proteinase
36	223	7.3	321	1 S27501	alkaline proteinase
37	219.5	7.2	613	2 S75976	hypothetical prote
38	218.5	7.1	272	2 A23624	subtilisin (EC 3.4
39	218.5	7.1	275	2 J01085	subtilisin (EC 3.4
40	218.5	7.1	533	1 J00146	serine proteinase
41	218.5	7.1	535	2 B82358	alkaline serine pr
42	218	7.1	361	2 G83756	subtilisin-type al
43	216.5	7.1	361	2 A48373	high-alkaline seri
44	216.5	7.1	525	2 G84406	haloalysin (impor
45	216.5	7.1	645	1 SUBSMP	serine proteinase

ALIGNMENTS

RESULT 1					
multidrug resistance transport protein - slime mold (Dictyostelium discoideum)					
C:Species: Dictyostelium discoideum					
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004					
C:Accession: T18279					
R:Shaulsky, G.; Loomis, W.F.					
submitted to the EMBL Data Library, June 1996					
A:Reference number: Z18855					
A:Accession: T18279					
A:Status: preliminary; translated from GB/EMBL/DBJ					
A:Molecule type: mRNA					
A:Residues: 1-1743 <SHA>					
A:Cross-references: UNIPROT:Q23868; UNIPARC:UP1000013687D; EMBL:U60086; NID:G1399914; P1					
A:Gene: tagc					
Query Match					
Best Local Similarity 24.2%; Pred. No. 3.8e-28;					
Matches 194; Conservative 114; Mismatches 213; Indels 281; Gaps 33;					
QY	68	HIHQFNGPIRETRKXL-----EXTGAKIIDYIPDYAYIV---EYEGDVXSK	131		
DB	115	YIVQFKRINDTRKQLKPLIGTDIVDQGYQSHIVHYIPHDSFLVMTQGVLSLS	174		
QY	132	XXXIHVESVEPY-----LPYXIDPOLFTKGASXLVK--AXALDT--KQXN	174		
DB	175	KEMVSWIGEPSPNNKIHLMVNEKSIQGLPVYII---LDSNMSLQRMENMLNSILKSYN	230		
QY	175	KEVQLRGI-----EXIAQXXXSNDVXYITAKPEYK	204		
DB	221	SKVKLTLLNQKQKSIYCNDESPSPCSLINSKLVYQWISQGESN---PIRSEKQ	287		
QY	205	VWMDVARGIV-----KADVAQSSYGLYGQGVYAVAVDTGLDTR---NNS----	246		
DB	288	TANRLSPKVFEGTKDTLVNNDKVDIP---LRGQILSLADTGLDSHCPSDSKPI	342		
QY	247	---SWHAPRGKITLVALGRTNNANDTNGGTHVAGSVLG-----NGXTNKGMAPQA	296		
DB	343	PLNSVNLNHR-KVVTYYITTSDDSDKVDGHTHICGSAAGPPESSVNISFSGLATDA	401		
QY	297	NLVFQSIQSGKGLGL--PSNLQTLFQOAYSAGARITNSWGA---AVNGAYTTDSRN	350		
DB	402	KIAF---FDLASGSSSLTPPSDLKQLYQPLVDAGARVHCDSSGVSVEGYTSSSDTAS	458		
QY	351	VDDYVRKR-DMTIIPAAQKXPNNGTIS--APGTAKMAYCATATENLR-----PSRES	400		
DB	459	IDDFLFTHPDFTILPAAGN---NEGYISLTLQSTAKNVITGAHQTIHENTYITDGPNYIN	515		
QY	401	YADNI-----	405		
DB	516	YQSSVDINQELICDDPSRYKCYTTTAQCCLESNATTGLASCPTLARKSVIDAANTQPILY	575		

QY 406 --NHVAQSSRGPTKQGRKIPDVMAFGTYILSARSSLA-----PDSSFWANHDSKAYVM 457
Db 576 NENNISSPSSKGPTRDKMKPALVAFGEYITSARNGANTTDQCCDGSF-PVTNALLA-I 633
QY 458 GGTSMATVIVAGNVNAQLREH-----FYKNGITPKPSLLKRALIAGA----- 499
Db 634 SGTSMATVIVAGNVNAQLREH-----FYKNGITPKPSLLKRALIAGA----- 499
QY 500 --ADGGLGYPNGN-----QGMGRVTLDKSLNVAAYNESS----- 531
Db 694 LITSSSIYPSNVQFENFAGASLVQGMGALIRMSNMLHYVNNNNNNNNKTSIDGITKFDGI 753
QY 532 -----XISTSQKATYXFT-----ATNGKPKL--ISLYWSDAPASTT 565
Db 754 GGLDLRLVYKPNQWKESSLSTGQNTSYCFYKPSSSSSNSGNNIPRVVATLWTDPPSYAG 813
QY 566 ASVTLVNDLVI-----TAPNGTYVGNDFPXXKXNMWDRNNVENVFINKP 613
Db 814 AKFNLVNNLDMITMYRRONGSTIFYSNOCSSFLG-----LAPQDTLNNVEGIVHNPT 867
QY 614 QSGTYTIEVQAYNPVGPQXFS 635
Db 868 EPMTRYFVAVAGTNVMPGFQNF 869

RESULT 2
T18267
multidrug resistance protein - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18267
R:Shaulsky, G.; Kuspa, A.; Loomis, W.F.
submitted to the EMBL Data Library, January 1995
A:Description: An MDR transporter/serine protease gene is required for prestalk specia1
A:Reference number: Z18850
A:Accession: T18267
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1905 <SHA>
A:Cross-references: UNIPROT:P54683; UNIPARC:UPI000013687C; EMBL:U20432; NID:g664839; PID
C:Genetics:
A:Gene: tagB

Query Match 15.6%; Score 477.5; DB 2; Length 1905;
Best Local Similarity 24.5%; Pred. No. 3.2e-25;
Matches 193; Conservative 98; Mismatches 237; Indels 259; Gaps 31;

QY 88 HIXQFNPGTLEETQXLE---XTGAKILDYIPDVAIVEX-----EGDYXS 130
Db 172 YIVQFKRINDREQLKEPLIGDITTLKEQPFKSHIVHYIPHDSFLVMTKQSVLLS 231
QY 131 XXXXIEHVESVEPY-----LPXYXI-----DPOLFTKASXL-----V 163
Db 232 SKEMISWIGHEPSSKIHNLNHEKSIQVPIYILSGSTNSLQRMENLNSILTSYNSKV 291
QY 164 KAXALDTQK-----XNKEVQLRGLEIXIAQXXKS--NDVYITTAPEYKVM 206
Db 292 KLTLLINQKLSIYCNDESPSSSSSSCSLIGSEKIYKWIISQESSENYIERSEKQTA 351
QY 207 NDVARGIV-----KADVAGSSYGLGGQGIYVADTGLDTGR--NDS----- 246
Db 352 NRLEPTVYFGTKDKLVNNDRIDP-----LRKGGLISLADTGLDGHCHPFSDSKTYPIPF 406
QY 247 -SMHEAFRGKITLALYALGRTNNANDTNGHGTIVAGSVLNGXTN-----KGMAPQANL 298
Db 407 NVQVNEHNRKVT---YIYHNEDEYVNGHGVHGSAACTPEDSSMAISPSGLATPAKI 463
QY 299 VFGSINSGXGLGULPSNLQTLFSGAKSAGARHTNSWGA---AVNGAYTTDSRNVDDY 354
Db 464 AFYD-LSSGSSEPTPEPDYSQMYKPLVDAGARVHGDWSGVSLSQYGYGDDAGGIDAF 522
QY 355 VRK-NDWTILFAAGNEXPGGTISAPGTAKNAITVGAENLRPRGS-----YADNI- 405

Db 523 LYEYPERSLPAAGNN-ELFASLLAQATAKNAITVGAELQTAHVNVSDALRYDESDAN 581
QY 406 -----NHVAQF 411
Db 582 FQRPCLPDKKVCNTTAKCCSESVNVKGLQCCPASIKQNASDFFTPQPPQYNNMGSF 641
QY 412 SSRGPTKQGRKIPDVMAFGTYILSARSS-----LAPDSFWANHDSKAYVMGTSMATP 465
Db 642 SSKGPTHDGRKLPDIVAFGEYITSARNGENSTDOCQGGSI--FNANGLMSISGTSMATP 699
QY 466 IVAGNVNAQLREH-----YKNGITPKPSLLKRALI----- 496
Db 700 LATATTLIKQYLVDFGFPYGESVVENKLPFTGSLKALMNNNAQLNGTYFWSNASSTNP 759
QY 497 AGAADXGLGYPNGQGMGRVTLDKSLNVAAYNESS-----XLISTQKAT----- 540
Db 760 SNAIFEQINGANLQGMGALFMN---NMLYKSSNPTPPSRWIGIGLGRKQKATWEMKD 816
QY 541 -----YXFT-----ATNGKP-LKISLYWSDAPASTTASVTLVNDLVI-- 576
Db 817 SLSSGLNKSXCFYKPSSSSSSGSGGCTPRIVATLWTDPPSYGAKFNLVNNLIDLILL 876
QY 577 -----VTTAPN--GTXYVGNDFPXXKXNMWDRNNVENVFINKPQSGTYTIEVQAYNP 628
Db 877 NSDDSIITIGNSGSLQPAKTAQF-----DTLNNVEGIIITPKANNYFTIAGTNVP 931

QY 629 VGPQXFS 635
Db 932 IGPKFS 938

RESULT 3
B83891
intracellular alkaline serine proteinase aprX [imported] - Bacillus halodurans (strain C
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: B83891
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saeaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: B83891
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <STO>
A:Cross-references: UNIPROT:Q9KBJ7; UNIPARC:UPI000003D43; GB:AF001513; GB:BA000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Gene: aprX

Query Match 11.5%; Score 350.5; DB 2; Length 444;
Best Local Similarity 30.2%; Pred. No. 3.9e-17;
Matches 114; Conservative 54; Mismatches 127; Indels 83; Gaps 17;

QY 183 EXIAQXXSNDVXITTKPKRYKNDVARGIVKA-DVAQSSYGLYGGQGIYVADTGLDT 241
Db 100 ESLQEMLVCKDIRKITYLNREYHALDTRAVESAQAFEYIRNGETLTGDOVIAVIDGT-- 157
QY 242 GRNDSMHEAFRGKITALY-ALGRTNNANDTNGHGTIVAGSVLNGXTN-----KGMAPQ 296
Db 158 -----YFHEDLEGRIKAFVDFVNRQREPYDNGHGTICADGMAAGNASSDQGTGRPAEA 212
QY 297 NLVFGSINSGXGLGULPSNLQTLFSGAKSAGARHTNSWGAAYVNGAYTTDSRNV----- 351
Db 213 NVIGIVKVLNQ--GWSGLESIHQV-----EMCIQYNBEHPDDPIHIISM 257
QY 352 -----DDYVF-----KNDWTILFAAGNEXPGGTISAPGTAKNAITVGAETEN 393
Db 258 GQALPYENBOEPDWVRIVEANNAAGITVCAAGNSGPDQQTASPSVSEKRVITVGLDD 317
QY 394 LRPSFGYANINNVAAQSSRGPTKQGRKIPDVMAFGTYILSARSSLAPDSF-----WA 448

Db 318 -RDYTDREDD---VAPFSRGPTLYGKPKPDILAPGVNIISLRS---PNSFYDKIOKGS 370
QY 449 NNDKXAYMGSTMAPPIYAGNAQAREHFVKKRGITTPRSLKALALIGA---AD-----501
Db 371 RVGSHYTMGSGTSMATPVCGVVALMLQH---BPNTLPDE--VKTILMGTDHMDRDPN 425
QY 502 -XGLGY-----PENGQ 511
Db 426 VYGAGYISABGAIPNSEE 443

RESULT 4

A69587

Intracellular alkaline serine proteinase aprX - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #ext_change 09-Jul-2004

C:Accession: A69587

R:Kumai, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beren

C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Brington, J.; Fabret, C.; Ferrati, R.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fume, S.; Galizzi, A.; Gall

A.; Jech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.

Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, R.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,

A.; Schlot, S.; Schroeter, R.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Seron

A.; Schlot, M.; Tanaka, A.; Tanaka, T.; Treppe, P.; Tognoni, A.; Tostato, V.; Uchiyama,

T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A.; Authors: Yoshikawa, H.F.; Zumschein, B.; Yoshikawa, H.; Darchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MID:198044033; PMID:9384377

A:Accession: A69587

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-442 <KUN>

A:Cross-references: UNIPROT:O31788; UNIPARC:UPI000006043F; GB:299113; GB:AL009126; NID:9

A:Experimental source: strain 168

C:Genetics:

A:Gene: aprX

F:146-398/Domain: subtilisin homology <SBT>

Query Match 10.4%; Score 319.5; DB 2; Length 442;
Best Local Similarity 30.0%; Pred. No. 5.8e-15;
Matches 106; Conservative 46; Mismatches 108; Indels 93; Gaps 15;

QY 202 EYKAWNDVARGIVKA-DVAQSYGLYGQGIYAVADYTGDRNDSSMHEAFRGKITLY 260
Db 120 EYKALDRTTBSHAKVAVRANGCITLTKVYAVVDYGI-----YHPDLGRI-----168
QY 261 ALGRTNNAN-----DTNGHGHVAGSYLNGXTN-----KGNAPQANLVFQSIMSKGL 310
Db 169 -IGPADMVNQKTEPYDDNGHGHVAGSYLNGXTN-----KGNAPQANLVFQSIMSKGL 226
QY 311 GGLPBNLQTLFQXASAGARHITNSWGAIVN-----GAYTTDSRNVDD 353
Db 227 GLTLADIRGV-----EMCIQYNDNDPDEPIDIMSMISGLDLYHQBQBP 272
QY 354 YVRKND-----MTILFAGNEXPNGGTISAPGTAKNAITVGATENLRPSFSYADINNH 407
Db 273 LVRAVEAKSAGIYCVANAGNSGSPDSQITIASPVSBEKVTYVGLADNDNTA-----SSDDT 328
QY 408 VAQFSRGPTKDGRIKPDVMAFGTILSARSSILAPDSSP-----NANHDSKAYMGSTSM 462
Db 329 VASFSRGPTVYGVKPKDILAPGVNIISLRS---PNSYIDKIQKSRVSGQFTMSGISM 385
QY 463 ATPYAGVNAQAREHFVKKRGITTPK--PELLKALALIGAADYGLYENGNGQW 513
Db 386 ATPICAGIALILQ---QNPDLTPDEVKELK-----NCTDKK 420

RESULT 5

T28159
pyrolysin (EC 3.4.-.-) - Pyrococcus furiosus
C:Species: Pyrococcus furiosus
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jul-2004
C:Accession: T28159
R:Vochozka, W.G.B.; Eggen, R.I.L.; Geerling, A.C.M.; Plateau, C.; Slezan, R.J.; Vos,
J. Biol. Chem. 271, 20426-20431, 1996
A>Title: Isolation and characterization of the hyperthermostable serine protease, pyroly

A:Reference number: Z20481; MID:9655370; PMID:8702780

A:Accession: T28159

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1398 <VOO>

A:Cross-references: UNIPROT:P72186; UNIPARC:UPI000016FDS9; EMBL:U55835; NID:G1556462; PI

A:Experimental source: DSM3638

C:Genetics:

A:Gene: p18

C:Keywords: hydrolase; serine proteinase

Query Match 9.7%; Score 297.5; DB 2; Length 1398;
Best Local Similarity 22.9%; Pred. No. 9.8e-13;
Matches 173; Conservative 76; Mismatches 248; Indels 257; Gaps 31;

QY 97 LEETKQXLEXTGAKTLD-----YIPDYATVYR-----GDYKXXXXIX 136
Db 77 LEETKQXLEXTGAKTLD-----YIPDYATVYR-----GDYKXXXXIX 136
QY 137 -HVESVEPYLPYXIDPOLFT-----KGSXLYKAXALDTKONKEV 177
Db 137 KQVTKRP-----SLBPKNYSTWVYVNLQFQEGYDQSGVAVVLDYGVDPNPFSLI 191
QY 178 QLRGLEXIAQKXKXNDVYITTKPEY-KYMN-----DYARGI-----213
Db 192 TPDGRKRIEWDFTDEGVDTSPFSKVNGLTINFTFYQVAGSLTNSGLMEYVK 251
QY 214 -----YKADVAGSYGLY-----GOGQIYVADT 237
Db 252 TYVSNVNTIGNTTSANGIYHFGLEPERFDLFPDQDQDFYVLLVNSGNGDYIAYDT 311
QY 238 GLDTGRNDS-----SMHEAFRGKITLYALGRTNNAN-----DTNGHGHVAG 280
Db 312 DLDYDFTBEPVLGQYNTVYDVAVPSYTGPLVYLAELIDPGEVAVFPGDGHGHVAG 371
QY 281 SYLNGXTN-----KGNAPQANLVFQSIMSKGL 308
Db 372 TVAGDSNNMDWMLSMYSGEWVPSRLYGMVDTYNTDTYQVAPGAGQIMAIRVLS-D 430
QY 309 GLGGLPSNLQTLFQXASAGARHITNSWGAIVNGLYT--TDSRN--VDDYVRKNDMTILF 364
Db 431 GKGSMWDLIRGN-TYAAHTGADVLSMSLQG--NADPYLDGTDPESVAVDELTEKGVVFT 487
QY 365 AAGNEXPNGGTISAPGTAKNAITVGATENLRPSFSYAD-----NIN 406
Db 488 AAGNEXPNGGTISAPGTAKNAITVGATENLRPSFSYAD-----NIN 406
QY 407 HVAQFSRGPTKDGRIKPDVMAFGTILSARSSILAPDSSP-----NANHDSKAYMGSTSM 466
Db 546 RIAPFSRGPTIDGRIKPDVMAFGTILSARSSILAPDSSP-----NANHDSKAYMGSTSM 466
QY 467 VAGNVAQAREHFVKKRGITTPKPSLKLALIGA-----DXGLYENGNGQW 513
Db 546 VAGNVAQAREHFVKKRGITTPKPSLKLALIGA-----DXGLYENGNGQW 513
QY 521 SLNVAVYNSKSLSTQKATYFPTATKPKLISLWSDAPASTA---SVTLVNDLVL 577
Db 655 SWEI-----LAINGTITLPIVDHMAKSYSDPAATLGLVIRGLYAR 696
QY 578 ITAPN-----GTYVGN-----DPXKPKXXNW-----DG-----RNVE 606
Db 697 NSIPDIVEMHKYVDTEYRTFEIYATBPWIKPFVSGVILNNTBFLARVYKVDGLBP 756
QY 607 -----NVFINXKPGSTYITIEVQANNVPVGPQXFS 635

Db 246 NMSLSSGSSSLIT--NAVDAYDKGVLIILAAAGSGKPPSGVFGALVNAVVALE 302
Qy 339 NLRPFGSADNINVAOSSRGPTDG-----RIKPDVMAFGTYILSARSLAPDSF 446
Db 303 NTON-GTY-----RAVDFSSRGKKTACDYYIQQKDVISAFAVAVST----- 346
Qy 447 WANHDSKYVMGTSNATPIVAGNVAOL 474
Db 347 W--PDGVTATISGTSMAFPAAGLANKI 372

RESULT 8

G83753
subtilisin-type protease (EC 3.4.21.-) vpr precursor [similarity] - Bacillus halodurans
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: G83753
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: AB3650; MUID:20512582; PMID:11058132
A/Accession: G83753
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-799 <STO>
A/Cross-references: UNIPROT:Q9KEM1; UNIPARC:UPI00000C39DC; GB:AF001510; GB:BA000004; NID
A/Experimental source: strain C-125
A/Genetics:
A/Gene: vpr
C/Superfamily: microbial serine proteinase vpr; subtilisin homology
C/Keywords: hydrolase; serine proteinase
F/1-29/Domain: signal sequence #status predicted <SIG>

Query Match 9.0%; Score 276.5; DB 2; Length 799;
Best Local Similarity 23.5%; Pred. No. 1.4e-11;
Matches 149; Conservative 61; Mismatches 198; Indels 227; Gaps 25;

Qy 78 LETVANNKMTXQNGP-TLE-----ETKQLEKTKAKILDTIPYATVEYE 125
Db 50 LETV-----IVIEDPSIIIAKIQGQKSKHEKQARQSVIEQIDLVF--SSTVTH 100
Qy 126 GDVAXXXXXIIEHVSVPFLPYKXIDPOLFTGSAKXIVAKXLDTRKQNKQVLAGIETI 185
Db 101 YDFLFGSALB-----LPAHQIPSLIGDGVHAVVPMIBEVETPDEY----- 144
Qy 186 AQXXXXNDVXYITAPRYK-VANDVARGIVKADVAQSSYGLYGQGIQVAVADTGLDTGN 244
Db 145 -----VIEKDAISPEMLDSAPFGANDAMEAGY--TGEGITVAIIDTGVDTYHP 191
Qy 245 DSSMHEAFRGKITLALYALGRTNNANDT-----NGHGTIVAGSVLNGXTKGMAPQ 295
Db 192 D-LVHAE-GDYKGMDFIDNDDPQETPPGDRGIBTHGTHVAGTVANGLI-KGVA PD 247
Qy 296 ANLYPQGISMDXSGIGSPSLQTL--FSQAKSAGARHTMSWGAIVAGATYDTSRNVDD 353
Db 248 ANLLAYRYL-----GGGSGSTAGVAGIERAQQDADIMNLSLGTLLNDPFAATIAL-D 302
Qy 354 YVRKNDMTILPAAGEXPNGGTISAPGAKAIVGATENLRPSGSY-----ADNINH 407
Db 303 WAMAEGVAVTSSNGSGPNNMTVSGPSTSRDAISVGAAT--RLPYNKTKASVFTSDGIDY 359
Qy 408 ----- 407
Db 360 PSADIMGFPSDEBELLELDEETRYEYAFAGLGRPDGFEVGVDEGKIALIVRGIFVEKXEN 419
Qy 408 ----- 407
Db 420 AKAGAVGAILYNNVAGVQPTVPGAIPITIMLSNEDGLKMRNELNGONTVTFISIEPKL 479
Qy 408 -----VAQSSRGPT-KQGRIKPDVMAFGTYILSARSLAPDSFPAHNSDKYVMGTSM 462
Db 480 VGERIVADSSSKQPVNHTMIKPDVSAFGVAVISTIPHQDDPY-----GYSGQGTSM 533

Qy 463 APPIVAGNVAOLREHFVNKRGITPKPSILKALINGAA-----DXGLGYDNGNGCRVTL 518
Db 534 ASPIVAGAAALILEKH-PWGV---DHVKALMTAENTLVDENGNRPHTQAG----- 584
Qy 519 DKSILNAVYVNESSXISTSOAKATY-XFTATAGKPLK 552
Db 585 --SIRIVDAISETLVTGSHSFGFTTKERKQVE 617

RESULT 9

S23407
subtilisin (EC 3.4.21.62) 1 precursor - Bacillus sp. (strain TA39)
C/Species: Bacillus sp.
C/Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 05-Oct-2004
C/Accession: S23407
R/Narinx, E.; Davail, S.; Feller, G.; Gerday, C.
Biochim. Biophys. Acta 1131, 111-113, 1992
A/Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarctic
A/Reference number: S23407; MUID:92256481; PMID:1581352
A/Accession: S23407
A/Molecule type: DNA
A/Residues: 1-420 <NAB>
A/Cross-references: UNIPROT:P28842; UNIPARC:UPI0000136183; EMBL:X62369; NID:G40200; PIDN
A/Genetics:
A/Gene: sub1
C/Superfamily: Subtilisin, subtilisin homology
C/Keywords: extracellular protein; hydrolase; serine proteinase
F/1-23/Domain: signal sequence #status predicted <SIG>
F/24-111/Domain: propeptide #status predicted <PRO>
F/112-420/Product: microbial serine proteinase #status predicted <MAT>
F/136-374/Domain: subtilisin homology <SBT>
F/145,185,360/Active site: Asp, His, Ser #status predicted

Query Match 9.0%; Score 274.5; DB 1; Length 420;
Best Local Similarity 28.3%; Pred. No. 7.9e-12;
Matches 107; Conservative 56; Mismatches 142; Indels 73; Gaps 21;

Qy 168 LDTKQXN--KEVQLRGIEIXIAQXXXNDVXYITAPK--YKVM--NDVARGIV---KA 216
Db 73 MNEKQFNLKKNKNTLVKVPKL-----RIATATDFEALVYNAAMASQTPMGIKATVNN 128
Qy 217 DVAGSYGLYQGGQVAVADTGLDTGRNDSMHEAFRGKITAL-YALGRT--NNANDT 271
Db 129 SITQTS---GGGGINIALDITGVNTN-----HPLRNNVQCKQFTVGTYYTNNSCTDR 179
Qy 272 NGHGTIVAGSVLNGXTNK-----GMAPQANLVFQSIM--DSXGGLGGLPSNLQTLFSQXS 326
Db 180 QHGTIVAGSALADGCTGNGYGVAPDADLMAVYKVLGDDGSGYADDIAMALIHAGDQATA 239
Qy 327 AGARITHTN-SWGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKYA 385
Db 240 LNTKVVIMSLGSSGESSLITNAVY---YSNKGVLILAAAGNSGPGYSGIYFPAALVNA 296
Qy 386 ITVGATENLRPSFGSYADNINVAQFSSRGPT-KDG-----RIKPDVMAFGTYILSARSS 439
Db 297 VAVVALEN-KYENGTY-----RVADFSSRGYSMTDGYAIOKGDVEISAFAAIVST--- 347
Qy 440 LAPDSSFPANNSDKAYVMGTSNATPIVAGNVAOLREHFVNKRGITPKPSI-----L 491
Db 348 -----W-FDGGATISGTSMAFPAAGLAAKIMQYPSASNDVAGELQYRAYENDI 398
Qy 492 KAALTAGAAD---XGLGY 506
Db 399 LSGYVAGYGDGDFASGFGP 416

RESULT 10

S11890
serine proteinase (EC 3.4.21.-) precursor, extracellular - Xanthomonas campestris pv. ca.
N/Alternate names: subtilisin-related proteinase
C/Species: Xanthomonas campestris pv. campestris
C/Date: 21-Nov-1993 #sequence_revision 07-Feb-1997 #text_change 31-Dec-2004

C:Accession: S11890
R:Lin, Y.N.; Tang, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
Mol. Gen. Genet. 220, 433-440, 1990
A:Title: A multipurpose broad host range cloning vector and its use to characterise an *E. coli* serine protease
A:Reference number: S11890; MUID:90251253; PMID:2187155
A:Accession: S11890
A:Molecule type: DNA
A:Residues: 1-580 <LTI>
A:Cross-references: UNIPROT:P23314; UNIPARC:UP000012A398; EMBL:X51635; NID:g48533; PIDN
A:Experimental source: Xanthomonas campestris pv. campestris
A:Note: The sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-ALA
C:Superfamily: extracellular alkaline serine protease precursor, Alteromonas type
C:Keywords: extracellular protein; hydrolase; serine protease
F:1-32/Domain: signal sequence #status predicted <SIG>
F:168-423/Domain: subtilisin homology <SBT>

Query Match 8.8%; Score 270; DB 2; Length 580;
Best Local Similarity 25.0%; Pred. No. 2.5e-11;
Matches 135; Conservative 66; Mismatches 194; Indels 146; Gaps 26;

QY 162 LVKA-XALDTKXNKEVQLRGEXIAQXXXSNDVXYITAKP-EYKVMNDVARGIVKA--D 217
DB 99 LVRAALDRABAEFLMRQLAADPNVGVENVQILHATLTLPNDTRLSEQMAFGTTNAGLN 158
QY 218 VAQSSYGLYGCGQIVAVADTG---DTGRDSSMHEAFRKGITLALVAGRTNAND--- 270
DB 159 IRPAMDRAVSGGVAVVADTGTSHADLNANILAGYDPTSDATVARDGNSDMSADGCD 218
QY 271 -----TNGHGVHAGSV--LGNGKTN-KGMAPOANLVPQSLMDSXG--- 309
DB 219 WYANECGAGIPAASSMHGTHVAGTVAATNNTTGVAGTVAQVVRVLGKCGSL 278
QY 310 -----LGGLPSNLQ--TLFQOASAGARHITNSGAAVNGAYTTDSRNVDY 354
DB 279 DIADALVWASGQVSGIPANANPAEVINMSLGCGSCSTTQON-AINGAVSRGT----- 331
QY 355 VRKNDMTLPAAGNEXPNVGGTISAPGTAKNAITVATEN--LRPSFGSYADNINNVAFS 412
DB 332 -----TVVAAGNDASVNSG-SLPANCANVAVAAVTSAGAKASVSPFGTGI----- 377
QY 413 SRGPKDRIKPDVWAPGTXTLSARS--LAPDSFMANHDSKYVMGTSMAFPIVAGN 470
DB 378 -----DVSAPESSILSTLNSGTTTSGS-----ASYASVNGTMSAPHVAGV 418
QY 471 VAQREHFVKNRGITPK--PSLTK--AALLAGADXLGY-----PNCNQ 511
DB 419 VALVQS--VAPYALTPAVERETLKNTARALPGACGCGAGIVNADAATVAINGSGSG 476
QY 512 GVGWVTLDKSLNVAVYNSKSLSTGOKATYKFTATAGKPKLSLVWSDAPASTTASVTL- 570
DB 477 GGGGNTLTNGTPVTGLG---AATGAELNYTITVPAG-----SGTITVTT 518
QY 571 --VNDLDELVI---TAPNGTYVGNDFXXPKXXNMGRNVENVFINKPQSGTYTIEVOY 625
DB 519 GSGSDADLYVAGSAPTSATYCRPYRS-----GNAEFCITTA-SGTYVRLKAY 568
QY 626 N 626
DB 569 S 569

RESULT 11
JC4908
alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.
N:Alternate names: subtilase
C:Species: Alteromonas sp.
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 05-Oct-2004
C:Accession: JC4908
R:Tsujido, H.; Miyamoto, K.; Tanaka, K.; Kaida, Y.; Inada, C.; Okami, Y.; Inamori, Y.
Biochem. Biotechnol. Biochem. 60, 1284-1288, 1996
A:Title: Cloning and sequence analysis of a protease-encoding gene from the marine bacterium *Alteromonas*
A:Reference number: JC4908; MUID:97141200; PMID:9897544

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-715 <TSU>
A:Cross-references: UNIPROT:P70765; UNIPARC:UP000017A928; DDBJ:D38600; NID:g1536787; PI
A:Experimental source: strain O-7
A:Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-sensit
C:Genetics:
A:Gene: aprI
C:Superfamily: extracellular alkaline serine protease precursor, Alteromonas type
C:Keywords: hydrolase
F:1-40/Domain: signal sequence #status predicted <SIG>
F:41-150/Domain: amino-terminal propeptide #status predicted <ATP>
F:151-486/Product: alkaline serine protease I #status predicted <MAT>
F:182-452/Domain: subtilisin homology <SBT>
F:497-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 8.8%; Score 268.5; DB 2; Length 715;
Best Local Similarity 25.8%; Pred. No. 4.3e-11;
Matches 142; Conservative 51; Mismatches 201; Indels 157; Gaps 28;

QY 169 DTKQXNKEVQLRGEXIAQXXXSNDVXYIT-----AKREYKMD-----VARGIV 214
DB 117 DKLSAKBAQAQFMBEVA---SGNVEYIEIDQLKPFATPNDRYNDQWHTYEAAAGI- 171
QY 215 KADVAQSSYGLYGCGQIVAVADTG---LDTGRDSSMHEAFRKGITLALVAGRTNAND 270
DB 172 NAPAAMDK--ATGGGVAAVLDGTGRPHLDLDAALITPGDWISNTFVANDGARDNDARD 229
QY 271 -----TNG-----HGHVAG---SVLGNGXTKMGAPQANLVPQS 302
DB 230 PGDAVTRGECGTDSGQVPRPADDSMHGTHVAGTVAATYNNGBGVAVYDAKVPVR 289
QY 303 IMDSXGGLGGLPSNLQTLFQOASAGARHITNSGAAV-----NGAYTTDSRNVDY 355
DB 290 VL---GKCGGLTSDIADGIWASGSDRVPANANPAVIVINMSLGCGACSAATONALNOA 346
QY 356 RKNDMTLPAAGNEXPNVGGTISAPGTAKNAITVAT--ENLRPSFGSYADNINNVAFS 413
DB 347 RNNGTIVIVAGNDNDNSANYN--PGNCGVNVAVASVGRDSRAYSYNGAI----- 397
QY 414 RGPDKDRIKPDVWAPGTXTLSARS--LAPDSFMANHDSKYVMGTSMAFPIVAGN 465
DB 398 -----DVAAFG---GAQSPADDEPGILSTNNSGSGAPNSYHSQTSMAAP 442
QY 466 IVAGNVAGLR-----EHFVN--RGITPKSLRAALI--AGADXLG---Y 506
DB 443 HVAGVAAALIKQAKSATPDEVETILKNTTRSPAGSCNCGGVDAALAAVVEALGADVTP 502
QY 507 PNGQSGRVTLDKSLNVAVYNSKSLSTGOKATYKFTATAGKPKLSLVWSDAPASTTASVTL- 566
DB 503 PTGN-----TLSD--GVAKTGLSGAAGSNQ--FTPDVDPAG-----TNV 538
QY 567 SVTL---VNDLDELVI---TAPNGTYVGNDFXXPKXXNMGRNVENVFINKPQSGTYTIEVOY 619
DB 539 TPTWSGCTGADLDLYK-----LGSG---PTSSSYCRPYEGGNALVCSDARQACTYH 588
QY 620 IEQAVNVVPG 630
DB 589 VMINGYKAYSG 599

RESULT 12
F69730
cell wall-associated protein precursor wpra - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: F69730
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berta
C.; Bron, S.; Broillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerzon, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier

leach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsberg, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapides, A.; Lardinois, A.; Lamber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Maue, Y.M.; Ogawa, K.; Ogiwara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli, Rieger, M.; Rivolta, C.; Roche, R.; Roche, B.; Rose, M.; Sadai, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schreier, R.; Scofield, P.; Seiguchi, J.; Sekowska, A.; Serot, A.; T. Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yoshikawa, H.P.; Zimstein, E.; Yoshikawa, H.; Danchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A/Reference number: A69580; MUID:9804403; PMID:9384377
A/Accession: P69730
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-894 <RUN>
A/Cross-references: UNIPROT:P54423; UNIPARC:UPI0000030283; GB:299109; GB:AL009126; NID:9
A/Experimental source: strain 168
C/Genetics:
A/Gene: wprA
F:453-664/Domain: subtilisin homology <SBT>

Query Match

8.6%; Score 262; DB 2; Length 894;
Best Local Similarity 23.9%; Pred. No. 1.7e-10;
Matches 134; Conservative 92; Mismatches 211; Indels 124; Gaps 28;

QY 50 PSKQOTGAAP--LVSENVKXKGLKKLRTVPA-----NNKLHIXFN---G 94
DB 245 PDVDSNGKSSSKYTERKQOKAINTLYDALQSVFSLKEIKQADRLMKQLQCKITAG 304
QY 95 PILER---TKQKLETKGKILDIPIYAVIV---EYSG-----DVKSXXX 134
DB 305 ALITENNIAKSEVQGT--KVIFKVDNLSLSVNHENKGFSAQSKDISVVKAKKL 362
QY 135 IEHVESVEYLLPYKXIDPOLFTKGASLYKAAALDTQKQKVEYQLGIEIXIAQ-XXXSND 193
DB 363 PNLVSFE--LPDEKQNGAYTASAKRYVSAATLSKMSNVEP---AEPVQEKSLAND 416
QY 194 VYIYIAKPEYKVNDAVARGIVK-ADVA-QSSYGLYQGGQ-----IYVADTGLDTGDNDS 247
DB 417 IQY---PYQWPLKNGENGVKQADVKYEPANTLSKRLDNTLLAVDVTGVDSTAD-- 471
QY 248 MHBAFRGKITALYA--LGRTNANDTNGHGVNAGSVLG--NGXTNKGMAPOANTLVQ 301
DB 472 ---LKGKVRTDLGNHFNVRNNNANDQGHGTHVAGIIAAGDNGSMGTGLAKAKITPV 527
QY 302 SIMDSXGIGLPSNLQTLFPOAKSHGARIHTNSWGAANVAYTDSRVVDYVRKQNT 361
DB 528 KVLDSAG--SGDTEQIALGIRYAADKAGAVINLSIG---GGYSRLBFLAKYAADKNTL 581
QY 362 ILFAANEXPNNGTISAPGTAKRATVYGATENTLARPFGSYADNINIVAPSSRGPTKQOR 421
DB 582 IAAAGNDEN--AISTPASKYVMSVGT-----NRMMDTADFSYNGKL--- 625
QY 422 IKPDVAPGTXILSARSSILAPDSSFANHDSKYAVMGTSMTPIYAGVNAOLREHFVN 481
DB 626 ---DISAPGSDI-----PSLVNNGN-----VTNMSGTSMATPYAAAAAGLLFPAQNPGL 670
QY 482 KQITPESILKAA--LIAGADYG-----LQYP-----NGNQGVRVTLIDKSLN 523
DB 671 KRTVEEDMLKTYADDISFESVDSGEBELDYDGDPIEIKPTGVDMHSGYGLNWKAVS 730
QY 524 VA-----YVNSKSLSTOKAT 540
DB 731 MADLQKVKKLESTOTAVRGS 751

RESULT 13

729090
surface layer-associated STABLE protease - *Staphylothermus marinus*
N/Alternate names: hyperthermostable protease
C/Species: *Staphylothermus marinus*
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
C/Accession: T29090

R/Mayr, J.; Lupaş, A.; Kellermann, J.; Eckerskorn, C.; Baumeister, W.; Peters, J.
Curr. Biol. 6:739-749, 1996
A/Title: A hyperthermostable protease of the subtilisin family bound to the surface layer
A/Reference number: Z20559; MUID:96385442; PMID:8793300
A/Accession: T29090
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1345 <MAY>
A/Cross-references: UNIPROT:Q54437; UNIPARC:UPI0000062778; EMBL:U057968; NID:91374755; P
A/Experimental source: strain F1
C/Function:
A/Note: serine protease S-layer component

Query Match

8.5%; Score 260.5; DB 2; Length 1345;
Best Local Similarity 29.1%; Pred. No. 3.7e-10;
Matches 98; Conservative 41; Mismatches 105; Indels 93; Gaps 14;

QY 252 PRGKTALVALRTNANNNDNGHGVTA-----GSVL-----GNGTKN--GMAPQANLV 299
DB 445 YQGRYAL-----VDFHGHGSVATVIASRGKVLVDLYGDKLYRIMGVAPGAKI- 495
QY 300 POSINDSKGLGSLPSNLQTLFPOAKSAG-----ARITNSM 336
DB 496 -----ACGDAMLGNIIVL--EAMLAGFNIVTEREDGYVLSLDPGPHRADITSNG 546
QY 337 GAANVAYTDSRVND-----DYRKQDNTILFAAGNXPNGGTISAPGTAK 383
DB 547 GSIYINFWLQPFPGIDYRSSFMDEILAIHNYLIGHVITVFAAGNPGVSSNGAROTGL 606
QY 384 NAITGATE--NLRSFG--SYADNINIVAPSSRGPTKQRIKPDVAPG-----TX 432
DB 607 LVITAGASTLMDYTRLYGPEGYAD--EVIFFSSGPPGQGYPKPDYINIGAFEMASNR 663
QY 433 ILSARSSILAPDSSFANHDSKYAVMGTSMTPIYAGVNAOLREHFVNKRGTIPKPSLX 492
DB 664 TIDGRGYAQPDVF-----GRTSEALPTYSGLTALPQAYKEVYNTTDPDYAK 712
QY 493 AALAGADXLGYPNQNGMGRVTLIDKSLNAVYNE 529
DB 713 IILKSSAKD--IWPYAFSGSGRVDAKKAADYVIFSE 747

RESULT 14

A33973
high-alkaline serine protease (EC 3.4.21.-) yab precursor - *Bacillus* sp. (strain yab)
N/Alternate names: alkaline elastase; subtilisin yab
C/Species: *Bacillus* sp.
C/Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 05-Oct-2004
C/Accession: A33973; B33973
R/Kaneko, R.; Koyama, N.; Tsai, Y.C.; Uang, R.Y.; Yoda, K.; Yamaoka, M.
J. Bacteriol. 171, 5232-5236, 1989
A/Title: Molecular cloning of the structural gene for alkaline elastase yab, a new subtilisin-like serine protease from *Bacillus* sp.
A/Reference number: A33973; MUID:89359181; PMID:2670913
A/Accession: A33973
A/Molecule type: DNA
A/Residues: 1-378 <KAN>
A/Cross-references: UNIPARC:UPI0000129678; GB:M28537; NID:9341960; PID:AAA67324.1; PID:
A/Note: parts of this sequence, including the amino end of the mature protein, were determined from the cDNA sequence.
A/Accession: B33973
A/Molecule type: protein
A/Residues: 111-164/326-355 <KA2>
A/Cross-references: UNIPARC:UPI0000175C84; UNIPARC:UPI0000175C85
C/Genetics:
A/Gene: ale
A/Start codon: TTG
A/Superfamily: Subtilisin, subtilisin homology
C/Keywords: extracellular protein; hydrolase; serine protease; zymogen
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-110/Domain: activation peptide #status predicted <ACP>
F:111-378/Product: alkaline elastase #status predicted <MAT>
F:132-338/Domain: subtilisin homology <SBT>
F:141,171,324/Active site: Asp, His, Ser #status predicted

QY 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGKXTNKGMAPOANLVE 300
DB 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGKXTNKGMAPOANLVE 300
QY 301 QSIMDSXGGLGGLPSNLQTLFSGOAXSAGARIHTNSGAAVNGAYTTDSRNVDVYRKNDM 360
DB 301 QSIMDSXGGLGGLPSNLQTLFSGOAXSAGARIHTNSGAAVNGAYTTDSRNVDVYRKNDM 360
QY 361 TILPAAGNEXPNVGGITISAPGTAKNAITVCATENLRPSFGSYADNINHVAQSSRGPTKDG 420
DB 361 TILPAAGNEXPNVGGITISAPGTAKNAITVCATENLRPSFGSYADNINHVAQSSRGPTKDG 420
QY 421 RIKPVPVMAPGTITLSARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVQLREHFPK 480
DB 421 RIKPVPVMAPGTITLSARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVQLREHFPK 480
QY 481 NRGITPKPSLTKAALAGAADXGLGYPNGNGMGWVTLDKSLNVAAYNNESSXLTSGKAT 540
DB 481 NRGITPKPSLTKAALAGAADXGLGYPNGNGMGWVTLDKSLNVAAYNNESSXLTSGKAT 540
QY 541 YXFITAGKPLKISLWMSDAPASTTASVTLVNDLDTITAPNGTYVGNDFPKXPKXKXND 600
DB 541 YXFITAGKPLKISLWMSDAPASTTASVTLVNDLDTITAPNGTYVGNDFPKXPKXKXND 600
QY 601 GRNNENVFINKPOSGTYTIEVOAYNVPGPOXPSLAIYN 640
DB 601 GRNNENVFINKPOSGTYTIEVOAYNVPGPOXPSLAIYN 640

RESULT 2

Q9AOR3_9BACI PRELIMINARY; PRT; 639 AA.

AC Q9AOR3; 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Protease.
GN Name:PROA;
OS Bacillus sp. 9860.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxID=133778;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=9860;
RX MEDLINE=20568675; PubMed=1118284; DOI=10.1006/birc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046403; BAB21266.2; -; Genomic_DNA.
DR HSSP; P00782; 1SUP.
DR SMR; Q9AOR3; 206-639.
DR MEROPS; S08.123; -;
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilisin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR002029; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASB_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 639 AA; 68186 MW; 316AF6FDBE4FF54 CRC64;

Query Match 98.7%; Score 3020; DB 2; Length 639;
Best Local Similarity 93.4%; Pred. No. 1.4e-209;
Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 4 KKKVFLSVLSAAAIITSTVALXNPSAGXARFEDLPFKGIQTITDDXXGFSKQKQTGAALFLV 63
DB 3 KKKVFLSVLSAAAIITSTVALXNPSAGXARFEDLPFKGIQTITDDXXGFSKQKQTGAALFLV 62
QY 64 ESENVKLXKGLXKKLETVPANNKLIHQFNCPILIEETXQILEXTGAKILDIYIPVAYIVE 123
DB 63 ESENVKLXKGLXKKLETVPANNKLIHQFNCPILIEETXQILEXTGAKILDIYIPVAYIVE 122
QY 124 YEGDVXSKXXXIENHVESVEPLPKYKIDPOLFTYGASLTVKXALDTPQXNKXVQLRGIE 183
DB 123 YEGDVXSKXXXIENHVESVEPLPKYKIDPOLFTYGASLTVKXALDTPQXNKXVQLRGIE 182
QY 184 XIAQXXXNDVXYITAKPEYKAMDVARGIYKADVAQSSYGLYGGQGVAAVADTGLDGR 243
DB 183 EIAQTVASNDVHYITAKPEYKAMDVARGIYKADVAQSSYGLYGGQGVAAVADTGLDGR 242
QY 244 NDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGKXTNKGMAPOANLVEQSI 303
DB 243 NDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGKXTNKGMAPOANLVEQSI 302
QY 304 MDSXGGLGGLPSNLQTLFSGOAXSAGARIHTNSGAAVNGAYTTDSRNVDVYRKNDMTIL 363
DB 303 MDSXGGLGGLPSNLQTLFSGOAXSAGARIHTNSGAAVNGAYTTDSRNVDVYRKNDMTIL 362
QY 364 FAAGNEXPNVGGITISAPGTAKNAITVCATENLRPSFGSYADNINHVAQSSRGPTKGRIR 423
DB 363 FAAGNEXPNVGGITISAPGTAKNAITVCATENLRPSFGSYADNINHVAQSSRGPTKGRIR 422
QY 424 PDVMAPGTITLSARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVQLREHFPVNRG 483
DB 423 PDVMAPGTITLSARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVQLREHFPVNRG 482
QY 484 ITPRPSLTKAALAGAADXGLGYPNGNGMGWVTLDKSLNVAAYNNESSXLTSGKATYTF 543
DB 483 ITPRPSLTKAALAGAADXGLGYPNGNGMGWVTLDKSLNVAAYNNESSXLTSGKATYTF 542
QY 544 TATGKPLKISLWMSDAPASTTASVTLVNDLDTITAPNGTYVGNDFPKXPKXKXNDGRN 603
DB 543 TATGKPLKISLWMSDAPASTTASVTLVNDLDTITAPNGTYVGNDFPKXPKXKXNDGRN 602

RESULT 3

Q76L84_9BACI PRELIMINARY; PRT; 640 AA.

AC Q76L84; 05-JUN-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Protease.
OS Bacillus sp. KSM-9865.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxID=192495;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KSM-9865;
RA Okuda M., Saeki K., Kobayashi T.;
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB084155; BAC82522.1; -; Genomic_DNA.
DR HSSP; P00782; 1AON.
DR SMR; Q76L84; 207-640.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilisin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR002029; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KM Hydrolyase; Protease; Serine protease.
 SQ SEQUENCE 640 AA; 67940 MW; 42187A150F2868F CRC64;

Query Match 98.4%; Score 3010; DB 2; Length 640;
 Best Local Similarity 92.8%; Pred. No. 7.6e-209;
 Matches 594; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 MEXKKKVPISVYSAALISTVALXNPAGAKXFPDIDFGIQTITDDXKSPKQQTGAAA 60
 DB 1 MEKKKKVPLSVSAALISTVALSNPSAGANFPDIDFGIQTITDDXKSPKQQTGAAA 60
 QY 61 FLVESBNVLTXXGKLKQLETVPANNGLHXORNGPILSEKTKXLBXTGAKITDIYIPDYAV 120
 DB 61 FLVESBNVLTXXGKLKQLETVPANNGLHXORNGPILSEKTKXLBXTGAKITDIYIPDYAV 120
 QY 121 IVEYEGDVAXXXXXXIEHVESVEPYLPKXXIDPOLFTKGASXLVKAXALDTKXNKVEQLR 180
 DB 121 IVEYEGDVAXSATSTIEDVESVEPYLPYRIDPOLFTKGASXLVKAXALDTKXNKVEQLR 180
 QY 181 GIEKTAQXXXXXNDVXYITNAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIYAVADTGLD 240
 DB 181 GIEQIAQFATSNVDVYITNAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIYAVADTGLD 240
 QY 241 TGRNDSMHEAFRGKITLALYALGRTNADTNGHGHVAGSVLGNSTKGMAPQANLVF 300
 DB 241 TGRNDSMHEAFRGKITLALYALGRTNADTNGHGHVAGSVLGNSTKGMAPQANLVF 300
 QY 301 OSIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDM 360
 DB 301 OSIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDM 360
 QY 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
 DB 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
 QY 421 RIKPVPVAPGTIILARSLSLAPDSSFWANHDSKRYVMGTSMAPTPIVAGNVAOLEHFPVK 480
 DB 421 RIKPVPVAPGTIILARSLSLAPDSSFWANHDSKRYVMGTSMAPTPIVAGNVAOLEHFPVK 480
 QY 481 NRGITPKPSLKAALIAAGADXLGYPNNGQMGRTVLDKSLNVAVYNSSXLSSTOKAT 540
 DB 481 NRGITPKPSLKAALIAAGADXLGYPNNGQMGRTVLDKSLNVAVYNSSXLSSTOKAT 540
 QY 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLDTLITAPNGTYVGNDFPKXPKXND 600
 DB 541 YSFTATAGKPLKISLWSDAPASTTASVTLVNDLDTLITAPNGTYVGNDFPKXPKXND 600
 QY 601 GNNNENVPINXPOSGTYTIEVOAYNVPGPQFSLATVN 640
 DB 601 GNNNENVPINXPOSGTYTIEVOAYNVPGPQFSLATVN 640

RESULT 4
 Q9AQR0_9BACT PRELIMINARY; PRT; 434 AA.
 AC Q9AQR0;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Protease (Fragment).
 GN Name=PROB;
 OS Bacillus sp. NV1.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=133781;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=NV1;
 RX MEDLINE=2058675; PubMed=1118284; DOI=10.1006/birc.2000.3931;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 RA Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 alkaliophilic Bacillus spp.: enzymatic properties, sequences, and

RT evolutionary relationships.";
 RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
 DR EMBL; AB046406; BAB21265.1; -; Genomic_DNA.
 DR HSSP; P00782; ISUP.
 DR SMR; Q9AQR0; 1-434.
 DR MEROPS; S08.123; -.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilisin activity; IEA.
 DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR007280; Pept_Bact_C.
 DR InterPro; IPR000209; Pept_S8_S53.
 DR Pfam; PF004151; ppc; 1.
 DR Pfam; PF004151; ppc; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KM Hydrolyase; Protease; Serine protease.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 434 AA; 45294 MW; 83517EDDB74125D2 CRC64;

Query Match 68.1%; Score 2082; DB 2; Length 434;
 Best Local Similarity 91.5%; Pred. No. 5.2e-142;
 Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 207 NDVARGIKADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 266
 DB 1 NDVARGIKADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 266
 QY 267 NADNTNGHGHVAGSVLGNSTKGMAPQANLVPOSINDSXGGLGGLPSNLQTLFQOAXS 326
 DB 61 NADNTNGHGHVAGSVLGNSTKGMAPQANLVPOSINDSXGGLGGLPSNLQTLFQOAXS 326
 QY 327 AGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDMTILPAAGNEXPNGGTISAPGTAKNAI 386
 DB 121 AGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDMTILPAAGNEXPNGGTISAPGTAKNAI 386
 QY 387 TYGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKDVMAPTPIILARSLSLAPDSSF 446
 DB 181 TYGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKDVMAPTPIILARSLSLAPDSSF 446
 QY 447 WANHDSKRYVMGTSMAPTPIVAGNVAOLEHFPVKRGITPKPSLKAALIAAGADXLGY 506
 DB 241 WANHDSKRYVMGTSMAPTPIVAGNVAOLEHFPVKRGITPKPSLKAALIAAGADXLGY 506
 QY 507 PNGNGMGRTVLDKSLNVAVYNSSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 566
 DB 301 PNGNGMGRTVLDKSLNVAVYNSSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 566
 QY 567 SYTLVNDLDTLITAPNGTYVGNDFPKXPKXNDGRNENVPINXPOSGTYTIEVOAYN 626
 DB 361 SYTLVNDLDTLITAPNGTYVGNDFPKXPKXNDGRNENVPINXPOSGTYTIEVOAYN 626
 QY 627 VPVGPOFSLATVN 640
 DB 421 VPVGPOFSLATVN 434

RESULT 5
 Q9AQR1_9BACT PRELIMINARY; PRT; 433 AA.
 AC Q9AQR1;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Protease (Fragment).
 GN Name=PROD;
 OS Bacillus sp. SD521.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=133780;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SD521;

```

RX MEDLINE=2056675; PubMed=1118284; DOI=10.1006/bbrc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RT Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046405/BAB21268.1; -; Genomic_DNA.
DR HSSP; Q45670; IDBI.
DR SMR; Q9AQR4; 1-433.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR00209; Pept_Bact_C.
DR InterPro; IPR00280; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Hydrolase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45576 MW; 98A2DF18FE60DDC CRC64;

Query Match 63.8%; Score 1952.5; DB 2; Length 433;
Best Local Similarity 86.6%; Pred. No. 1.2e-132;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

QY 207 NDVARGIYKADVAAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 266
DB 1 NDVARGIYKADVAAQNNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITLALYALGRTN 60
QY 267 NANTNGHGTIVAGSVLNGKTNKGMAPOANLVFQSIIMDSXGGLGSLPSNLQTLFSAQXS 326
DB 61 NANTNGHGTIVAGSVLGN -ALNKGMAPOANLVFQSIIMDSXGGLGSLPSNLQTLFSAQMN 119
QY 337 AGARHTNSMGAANVGAATYDTSRNVDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKAI 386
DB 120 AGARHTNSMGAANVGAATYANSRQVDEYVRNNDMTVLPAGNBNPGTISAPGTAKAI 179
QY 387 TVGATENLRPSFGSYADNINHVAQFSRSGPTDGRIRKPDVMAFGTXLSARSSLPDSSF 446
DB 180 TVGATENLRPSFGSLADNPNIHQFSRSGATRDGIRKPDVTAFGTILSARSSLPDSSF 239
QY 447 WANHSKYAAMGGTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVGIGY 506
DB 240 WANYSKYAAMGGTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVGIGY 299
QY 507 PNGNGMGRVTLDSGLNVAAYNESSXLSQKATYXFTATAGKPLKISLVMSDAPASTTA 566
DB 300 PNGNGMGRVTLDSGLNVAAYNEATALTQKATYSFOQAKPKISLVMTDAGSTTA 359
QY 567 SVTLVNDLDLVITAPNGTXYVGNDFXPKXKNMGRNNVENVFINKPOSQTYTIEVQAYN 626
DB 360 SVTLVNDLDLVITAPNGKYVGNDFSYPDNMMGRNNVENVFINAPOSQTYTIEVQAYN 419
QY 627 VPVGPOKFSLAIVN 640
DB 420 VPVGPOKFSLAIVH 433

RESULT 6
Q9AQR4_9BACT PRELIMINARY; PRT; 433 AA.
AC Q9AQR4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Protease (Fragment).
GN Name=PROA;
OS Bacillus sp. D6.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
```

```

OX NCBI_TaxID=127889;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D6;
RX MEDLINE=2056675; PubMed=1118284; DOI=10.1006/bbrc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RT Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046402/BAB21265.1; -; Genomic_DNA.
DR HSSP; Q45670; IDBI.
DR SMR; Q9AQR4; 1-433.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR00209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Hydrolase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45636 MW; 52087E0A2516107F CRC64;

Query Match 63.7%; Score 1948.5; DB 2; Length 433;
Best Local Similarity 86.4%; Pred. No. 2.3e-132;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

QY 207 NDVARGIYKADVAAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 266
DB 1 NDVARGIYKADVAAQNNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITLALYALGRTN 60
QY 267 NANTNGHGTIVAGSVLNGKTNKGMAPOANLVFQSIIMDSXGGLGSLPSNLQTLFSAQXS 326
DB 61 NANTNGHGTIVAGSVLGN -ALNKGMAPOANLVFQSIIMDSXGGLGSLPSNLQTLFSAQMN 119
QY 337 AGARHTNSMGAANVGAATYDTSRNVDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKAI 386
DB 120 AGARHTNSMGAANVGAATYANSRQVDEYVRNNDMTVLPAGNBNPGTISAPGTAKAI 179
QY 387 TVGATENLRPSFGSYADNINHVAQFSRSGPTDGRIRKPDVMAFGTXLSARSSLPDSSF 446
DB 180 TVGATENLRPSFGSLADNPNIHQFSRSGATRDGIRKPDVTAFGTILSARSSLPDSSF 239
QY 447 WANHSKYAAMGGTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVGIGY 506
DB 240 WANYSKYAAMGGTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVGIGY 299
QY 507 PNGNGMGRVTLDSGLNVAAYNESSXLSQKATYXFTATAGKPLKISLVMSDAPASTTA 566
DB 300 PNGNGMGRVTLDSGLNVAAYNEATALTQKATYSFOQAKPKISLVMTDAGSTTA 359
QY 567 SVTLVNDLDLVITAPNGTXYVGNDFXPKXKNMGRNNVENVFINKPOSQTYTIEVQAYN 626
DB 360 SVTLVNDLDLVITAPNGKYVGNDFSYPDNMMGRNNVENVFINAPOSQTYTIEVQAYN 419
QY 627 VPVGPOKFSLAIVN 640
DB 420 VPVGPOKFSLAIVH 433

RESULT 7
Q9AQR2_9BACT PRELIMINARY; PRT; 433 AA.
AC Q9AQR2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
```

DE Protease (Fragment).
GN Name=PROC;
OS Bacillus sp. Y.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=133779;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley.
RX MEDLINE=20568675; PubMed=1118284; DOI=10.1006/dbrc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319 (2000).
DR EMBL; AB046404; BAB21267.1; -; Genomic_DNA.
DR HSSP; Q45670; 1DB1.
DR SMR; Q9AQR2; 1-433.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0004289; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_S8_S53.
DR InterPro; IPR007280; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASIS_HIS; 1.
DR PROSITE; PS00138; SUBTILASIS_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER
FT NON_TER
SQ SEQUENCE 433 AA, 45588 MW; B81291A803C775AE CRC64;

Query Match 63.5%; Score 1941.5; DB 2; Length 433;
Best Local Similarity 86.2%; Pred. No. 7.3e-132;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

QY 207 NDVARGIKADVAQSYGIGGQGVAVADTGLDGRDSSMHEAFRGKITLALYALGRTN 266
DB 1 NDVARGIKADVAQSYGIGGQGVAVADTGLDGRDSSMHEAFRGKITLALYALGRTN 60
QY 267 NADNNGHGTAVAGSVLNGXGKNGKMAPOANLVPQISINDSGGLGSLPSNLQTLPSQAXS 326
DB 61 NADNNGHGTAVAGSVLNGXGKNGKMAPOANLVPQISINDSGGLGSLPSNLQTLPSQAXS 119
QY 327 AGARIHTNSWGAANVAGAYTDSRNVDDYVRKNDMTILPAAGNEXPNGGTISAPGAKNAI 386
DB 120 AGARIHTNSWGAANVAGAYTDSRNVDDYVRKNDMTILPAAGNEXPNGGTISAPGAKNAI 179
QY 387 TVGATENTLRPSFGSADNININVAQSSRGPTKGRIKPDVMAPTXILSARSSLAAPDSGF 446
DB 180 TVGATENTLRPSFGSADNININVAQSSRGPTKGRIKPDVMAPTXILSARSSLAAPDSGF 239
QY 447 WANHSKYAVNGGSMATPIVAGNVAOLREHFVRKRGITPKPSLKALILAGAAAGXGIGY 506
DB 240 WANHSKYAVNGGSMATPIVAGNVAOLREHFVRKRGITPKPSLKALILAGAAAGXGIGY 299
QY 507 PNGNGGWRVTLIDKSLNVAAYVNESXLSSTQKATYXFTATAGKPKLISLVMSDAPASTTA 566
DB 300 PNGNGGWRVTLIDKSLNVAAYVNESXLSSTQKATYXFTATAGKPKLISLVMSDAPASTTA 359
QY 567 SVTLVMDLDVLTAPNGKTYGNDKFXPKXKXMDGRNVENVFVIXPOSGITTIIVQAVN 626
DB 360 SVTLVMDLDVLTAPNGKTYGNDKFXPKXKXMDGRNVENVFVIXPOSGITTIIVQAVN 419
QY 627 VPVGPQXPSLAIVN 640
DB 420 VPVGPQXPSLAIVN 433

RESULT 8
Q4NB18_9MICC
ID Q4NB18_9MICC PRELIMINARY; PRT: 697 AA.

AC Q4NB18;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Peptidase S8 and S53, subtilisin, kexin, sedolisin.
GN ORFNames=ArthDRAPT_0589;
OS Arthrobacter sp. FB24.
OC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxId=290399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24.
RG US DOE Joint Genome Institute (JGI-PCF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
Hammon N., Israni S., Pitluck S., Richardson P.;
RT Sequencing of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (PCF-ORNL);
RA Lartimer F., Land M.;
RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAHG0100025; EAL94539.1; -; Genomic_DNA.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASIS_HIS; 1.
DR PROSITE; PS00138; SUBTILASIS_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 697 AA; 73821 MW; B8924966C7C0714 CRC64;

Query Match 28.2%; Score 863.5; DB 2; Length 697;
Best Local Similarity 34.2%; Pred. No. 1.1e-53;
Matches 230; Conservative 98; Mismatches 215; Indels 129; Gaps 19;

QY 83 ANNKHLI-XQFNGPLLESTKQXLETKAKIIDYIPDAVYIVYEBQDVASXXXIIEHVSV 141
DB 30 ASESCHILIIQTRAPLTAQORAEIAGIDVEMQSYVDNTYLAAPADILRRVALPVSVA 89
QY 142 BPYLPRYXIIIDQLFTKGA-SKLVRKXALDTQKXNREYOL-----RGIEIXIAQXXX 190
DB 90 DYSRSVFRKIPPLPRASADTGNVRLADHEHPRRLERVDLLHPGLEAGPELIRAVNA 149
QY 191 S-----NDVXYITAKPEYKVMNDVARGIVADVAQ 220
DB 150 AARVPEDAVAVTGKLRITTSVGQLPEILAIDEIRHIVRBRQLFNNVAREILNANDVQL 209
QY 221 SSYGIYGGQGVAVADTGLDGRDSSMHEAFRGKITLALYALGRN--ANANDNNGHGTIV 278
DB 210 NGTTYRGAGEVAVADTGFDTG--DAANPAPAFGTGVQTLYALGRYAPKADDPHGHTIV 268
QY 279 AGSVLGNQXTN-----KGMAPQANLVPQISINDSGGLGSLPSNLQTLPSQAXSAGARIH 332
DB 269 AGSVLGNQXTN-----KGMAPQANLVPQISINDSGGLGSLPSNLQTLPSQAXSAGARIH 328
QY 333 TNSWGA-AVNGAYTDSRNVDDYVRKN-DMTILPAAGNEXPNG-----GTISAPGTA 382
DB 329 TNSWGPVGLATPYDASREIDEPVNNHDDQYICFAAGNDGVGNSDGTIVDSISGSAA 388
QY 383 KNATVGAETENLR-----SPGSY-----ADNININVAQSSRGPTKGRIK 423
DB 389 KNCTVGAESLRLKEFTPSYGTTPGDFPANPVKRDQKANNPDGVAVASSSRGPTKEGRK 448
QY 424 PDVNAFGTXILSARSSLAIP-DSSFVANHSKYAVNGGSMATPIVAGNVAOLREHFVRK 482
DB 449 PDVNAFGTXILSARSSLAIP-DSSFVANHSKYAVNGGSMATPIVAGNVAOLREHFVRK 508

```
Qy 483 GTPKPSLTKAALTAGA-----ADXLGPGNCGMGRVTLDKSLNV-----AYV 527
Db 509 LNSPAAALVKALLVAGADVLPQGINPSEAGBS-FNGNSGKRVNLARSVLTGQPGNAGL 567
Qy 528 NESSXLSTSOATYXXF-----TATGKPLKISLWSDAPAS 563
Db 568 GEGGDTGCGQSDSPFTIDIPBEVPKVAAGRRNRGPAALPAAGVTLKTLTWSDPRGP 627
Qy 564 TTASTTLYNDLDLVITAPNGIXYVGNDFKXPKXXNMDOGRNNVNFINKPQSGYTTIVQ 623
Db 628 -----QLNNDLDLVITLAADGSEHGN-----SGTTAGFDRNRNQLVMTGMPGQARIYR 679
Qy 624 AYNVPVGQXFS 635
Db 680 AFRITQPPQPYA 691

RESULT 9
Q4NVB5_9DELTA PRELIMINARY; PRT; 1748 AA.
ID Q4NVB5_9DELTA PRELIMINARY; PRT; 1748 AA.
AC Q4NVB5;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Peptidase S8 and S53, subtilisin, kexin, sedolisin;Nepovirus coat
DE protein, N-terminal;Nepovirus coat protein, N-terminal precursor.
GN ORName=AdenDRAFT_3007;
OS Anaeromyxobacter denalogenus 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cytophacteria; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Istant S., Pritch S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT denalogenus 2CP-C." to the EMBL/Genbank/DBJ databases.
RL Submitted (May-2005) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Laft M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT denalogenus 2CP-C." to the EMBL/Genbank/DBJ databases.
RL Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD0100011; EAL7953.1; -; Genomic_DNA.
KW Capsid protein; Signal.
FT SIGNAL 1 23
SQ SEQUENCE 1748 AA; 177411 MW; 118EDB338B81DE76 CRC64;

Query Match 25.9%; Score 793; DB 2; Length 1748;
Best Local Similarity 34.9%; Pred. No. 4.5e-48;
Matches 220; Conservative 86; Mismatches 217; Indels 108; Gaps 25;

Qy 89 IXOFNGFLERTKXKLEXTAKILDPYIVYEGDVXSKXXIXIHVSVEPYLKY 148
Db 103 IVQSGPFRTHKELAAAGARLDYDFEFAVMSDEATTAKYAMLFPVGAUKAEVY 162
Qy 149 XIDPOLF-TKGASXLV-----KAXALDTQXNKVEQLR 180
Db 163 KMAQQLADTSGAILEPGRSVRLVVDHDLAPVSLVBERKGMHLVARDASIRIL 222
Qy 181 GIEKIAQKXKSNVXYT-TAKREYKVMNDVARGIVK----ADVACSIGLYGQGIYVA 235
Db 223 DAD-IASLHLEVLWEPALPAY-VLNDTSRMTITQTVGSPDSTISDRGLGRGQIYAI 280
Qy 236 DTGAD-----TGRNDSMHEAFRGKITALVALGRTNNANDTN-GHGHVAGSVLGN 285
```

```
Db 281 DTGLDHACWFRDPLGAAAGPMR---KVAAYLTVC---GDDYDNLGHGHTVACTVAGD 334
Qy 286 -----GXTNKGAPQANLVFQSIIMDSXGLGGLPSNIQTLFQAKXSAGARIHTNSGAA 339
Db 335 QPTITGAAAGMAGPARGVVTVDLFLGNNMFSPPADLARIPTTPYALGARIHNTSMGSS 394
Qy 340 VNGAYTDSRVVDVYRKN-DMTILFAAGNEXPKGCTISAGTAKNALTVCATENTLRSF 398
Db 395 SN-AVDALARADRFMMHDPFLVLANGNAPVGVSGAATATKANVSVATGN----- 448
Qy 399 GSYADNINHVQFSSRGPTKDGRIKPDVMAEGTILSARSILAPDSSFWAHDSKYAYMG 458
Db 449 GLAABD---VASFSHGPAADGRKFTITAGCVGISADSIGTAS---NNGSVAF-S 500
Qy 459 GTSNATPIVAGNVAGLRREHYKN---RGI-----TPKSLTKAALTAG---AADXGL 504
Db 501 GTSNATPAAAGAAALVRRYFEGGFMPGLGSPADARSPSALVATLVNSAQNVVAGENG 560
Qy 505 G-YPNGNGGRVTLDSLVNA-----YNSSXSLSTOKATYXFTATGKPLKISLV 556
Db 561 GPISSTGGWGRINLSNLRPAADAAVLDVVEVAAGLETGGSFTROYSTGAQPLKTLV 620
Qy 557 WSDAPASTASVTLVNDLDLVITAPNG-TXYVGNDFX-----XPXXNMDOGRNNV 607
Db 621 WTDAPSGQLDRSLVNDLDLVITVPGGATYTLGNAVFLGSEVAGAP-----DLNVEQ 675
Qy 608 VFINKPQSGTYTIEVQAYNVPVGQXFSLA 638
Db 676 VLLAAPTGTGTVAVTGYNVPGQPFALVI 706
```

```
RESULT 10
Q6MKR4_BDEBA PRELIMINARY; PRT; 711 AA.
ID Q6MKR4_BDEBA PRELIMINARY; PRT; 711 AA.
AC Q6MKR4;
DT 05-JUN-2004 (TREMBLrel. 27, Created)
DT 05-JUN-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUN-2004 (TREMBLrel. 27, Last annotation update)
DE Subtilisin-like serine protease precursor.
GN OrderedLocusNames=Bd2321;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.D., Goemann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: 11fe cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective." to the EMBL/Genbank/DBJ databases.
RL Science 303:689-692(2004).
DR EMBL; BX842652; CAB80143.1; -; Genomic_DNA.
DR HSSP; P27693; 1AH2.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00138; SUBTILASIN_SER; 1.
KW Complete proteome; Protease; Signal.
FT SIGNAL 1 8
SQ SEQUENCE 711 AA; 74914 MW; 4930050843357E5 CRC64;

Query Match 25.7%; Score 786.5; DB 2; Length 711;
Best Local Similarity 30.9%; Pred. No. 4.1e-48;
Matches 229; Conservative 110; Mismatches 246; Indels 157; Gaps 25;

Qy 9 LSVLSAAILSTVALNPNRPAKXAPLDFKIGITTTDDXXGFSQXGTGAAPLVESNV 68
```

```

Db      8 ITSTVAALVLSVA-----AAGATVLKFNAG--AIDT-----NKLSTNSVAASMMMEAKAT 55
Qy      63 KAKKGLKXKLETVPANNKCHIXPNGBPILEETQKLEXTGAKLIDYIPDYAVYIV- YEGD 127
Db      56 E-----YIVQFKCAVTEKDKAALKAQ- FEYFGYLPDDALVVRGYSID 96
Qy      128 VASXXXXXIEHVESVEPYLPYXID-----PQLFTKASXKIY-----KAAA 167
Db      97 LVTFKTNHNGVQAVVKYAGNYKVSNSFAPASVFTKNNMAVLVNTFKSHEVEVIAKIEK 156
Qy      168 LPTKQKKEVQ-----LKGIEKI-----AQXXXSNDYXIT 198
Db      157 MGVKAVQVVDGKHITALIPRGIVPAVALVTVGEHVQPAVIESLHPADEDLADQYSAT 216
Qy      199 APEEYVMDVARG--IVKADVAQSSYGLYGQOIVAVDGLDTRGNDSSMHEAFRGKI 256
Db      217 AAGDYSIDLTVGDESGLTMNFDAAW-AMGYAGKQQTVMADTGLDSC-NGCALHQDPAGGV 274
Qy      257 TALVYALGRNNA-NDNNGHGVHAGSVLGNKXTNKM-----APQANLVFQSIIMDSXGIG 311
Db      275 IGGYFPGLSKSKSDPMGSGTHVAGSVMGRTASKGLKGAAYEAMVVAEGMWSPMWKQL 334
Qy      312 GLPSNIQTLFQKXKSAGARIHTNSMGA--VNGAYTTDSRVDDYVRKN-DMTILPAAGN 369
Db      335 SVPSKLGDFEKAFADGARIHTNSKCGARTFGAYDPAVQVDEMYANPDMILIPAAKNS 394
Qy      370 XP-----NGGTSAPGTAKNAITVGAATENI-----RPSFG 399
Db      395 GADKNKDGRIIDNSMSAPGTAKNVLTVGASENVTKSGGIQVPISKRAADKWPSPPIYS 454
Qy      400 SY-ADNINNVAPSSSPGPTKDGRIKPDVWAPGTXTLSARSSSLAPDSFPAANDSKAYNG 458
Db      455 STISDNGNLAMPSSSGPTTGDGKPDYVAPGTNVLSVPSQEKDAPLGAAYNKDVMWSG 514
Qy      459 GTSMAPIVAGNVAQLREHFVNKRGIT-TRKPSILKALJAGADKGLGY-----506
Db      515 GTSMAPIPLAAGAAALARQVLERKGMKNSALMKRTMLHTAVDMYRQGPGEIGARQGE 574
Qy      507 -----PNGNQMGKRVTLDKSLVA-----YVNESSXISTSOKATYXFTATAGKPLKISLVW 557
Db      575 ILTRPNSDEGGRVADVANIAMIGATQVDNRQGVAGQABVSYEFTLNAFGLVNLVW 634
Qy      558 SDAPSTTASVTLVNDLIVITAPNGTXVYVNDPKXPKXXKMDGRNNEVPIAKQSGT 617
Db      635 TDAPSANNAQALVNDLIVLTPNGQTLSPMDHI-----NNLEMIKSGLPACT 684
Qy      618 YTIETQAVVVPV---GPOKFSI 636
Db      685 YKLTIVKGPVPOGKNGAKAYAL 706

RESULT 11
0747P6 GEOSL PRELIMINARY; PRT; 2030 AA.
ID 0747P6 GEOSL PRELIMINARY; PRT; 2030 AA.
AC 0747P6-
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DB Fibronectin type III domain protein.
GN OrderedLucusName=GSUJ3219;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.B., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beaman M.J., Dodson R.J.,
RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Hatt D.H., Selengut J.,

```

```

RA Davidson T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R.,
RA Van Aken S.B., Loyley D.R., Fraser C.M.,
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT environments."
RL Science 302:1967-1969(2003).
DR EMBL: AE017180; AAR36610.1; -; Genomic_DNA.
DR HSSP: P27693; IAH2.
DR TIGR: GSUJ3219; -.
DR GO: GO:0004289; F:subtilase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR011635; APHP.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR000209; Pept_58_S53.
DR Pfam: PF07705; CARD; 8.
DR Pfam: PF00041; fn3; 1.
DR Pfam: PF00082; Peptidase S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
KW Complete proteome.
SQ
SEQUENCE 2030 AA; BD7ADC27DD141B01 CRC64;

Query Match      20.9%; Score 638; DB 2; Length 2030;
Best local similarity 29.4%; Pred. No. 8, 6e-37;
Matches 188; Conservative 93; Mismatches 243; Indels 116; Gaps 19;

Qy      83 ANKHLHXQPNPILEETQKLEXTGAKLIDYIPDYAVYIVYEGDYVXXXXXIEHVESVE 142
Db      54 AIRKRLVQFNQPNVPEGRQLEALGCRIGYMPFTAFAVLMDDKAAKVALLSFVEDIT 113
Qy      143 PLYPYXIIDPOLFTKASLYVKAALD-TRKQNKVEQLAGI-----182
Db      114 RFAF-----ADKLVTARKDLTAAPTSIRIKRVLRVDDPADRAAVIAATLKG 161
Qy      183 -----EXIAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAQS---221
Db      162 NGRILNVGARTTVEVPEBELAPLAQOETAMIGVGEIRLNSDAAMVQVTRNEVDNRIT 221
Qy      222 -SYGLYGQOIVAVDGLD-----TGNDSSMHEAFRGKITALYALGRTNANDT 271
Db      222 WEKGITGAGQIVIGIDSGVDYDMPFADPNGLPGGRHKIYG-----YDATLGDHDA 276
Qy      272 NGHGVHAGSVVGN---GKTNGMAPOANLVFQSIIMDSXGIGLPSNIQTLFQKXKSAG 328
Db      277 DGHGTHIGETICGDBGPGKNGIAPGARIHVDVGTGTLTG-SLBLETVLKRAYDSG 335
Qy      329 ARIHTNSMGAVNGAYTTDSRVDDYV-RKNDMTILPAAGNEXPNCGTISAPGTAKNAIT 387
Db      336 ARIHTNSMGAVNGAYTTDSRVDDYV-RKNDMTILPAAGNEXPNCGTISAPGTAKNAIT 387
Qy      388 VGATENLRPSPFGSYADNINNVAPQFSSRGPTKDGRIKPDVWAPGTXTLSARSSSLAPDSF 447
Db      395 VVATGN-----GTDAAI---VSAESSVGOAPDGRANPVGARPGQGVASRS---DGLIG 442
Qy      448 ANHDSKRYAMGTSMAPIVAGNVAQLREHF-----VKRNGITPKPSLLKALIA 497
Db      443 SENSCTMA-MSGTSVAALVTSGLAALIRQYDFDPFGTSPVATNKLPSALALLKAVLVN 501
Qy      498 GAA-----DXGLGVYNGNMGKRVTLDKSLV-----NVAYVNESSXISTSOKATYXFTA 545
Db      502 SAEALLSDPDGDSCKSGKGMGRPKLINTLFPNGSHSLVVVDGGTGLTGVWQRLYS 561
Qy      546 TAGKPLKISLVWSDAPSTTASVTLVNDLIVITAPNGTXVYVNDPKXPKXXKMDGRNNEV 599
Db      562 PGRRLKITTLAMTADPAAPAGATSPILTNDLNLVVAADGTTTIGNDLNSHGDYESRTGTF 621
Qy      600 DGRNNV-ENVPIAKQSGTITIEVQAVVVPVPGQKFSIAI 638
Db      622 SDRVNVBEQVYIKRPVAGTYVYKVGASIPVGPQFPALVM 661

RESULT 12
QAHUT5_GTBZE

```


ID 04HUY5 GIBZE PRELIMINARY; PRT; 1088 AA.
AC 04HUY5-
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG11223.1;
OS Gibberella zeae PH-1
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocryomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barua N., Bastien V., Bloom T., Boguslavsky L.,
RA Boudgaltier B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeRellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins G., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall U., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal A., Karatas A.,
RA Kellis C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-Y., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mäkelä E., McCarthy M., Meldrum J., Menais L.,
RA Mitova T., Mudge V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Nordu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachtupka A., Ramsamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schuer S., Schupbach R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Testaye S., Theodore J., Topham K., Travers M.,
RA Vasilev H., Venkataran V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence."
RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AACW01000460; EAA75433.1; -; Genomic_DNA.
CC Hypothetical protein.
SQ SEQUENCE 1088 AA; 119629 MW; E3B3CB94C07F542 CRC64;
KW
Query Match 19.4%; Score 594; DB 2; Length 1088;
Best Local Similarity 31.8%; Pred. No. 5.8e-34;
Matches 197; Conservative 98; Mismatches 231; Indels 94; Gaps 29;
QY 96 ILEETKQLEKTKGAKILDPYAYIVE-YEGDVYKXXXXIEHVSVEPYLPXYXIDPOL 154
DB 476 ILITQKQ---LTGSPYLAELKSTISLKMWDRELDRTYRDCILHEPNTYAEKLAPEI 532
QY 155 FTGKASXLVKKALDPTQKXNKEVQGLGIEKIAQXXXXNDVYTTAKPEKYMNVDVARGIV 214
DB 533 AEKAGVDITK-LAVSPVRIRLTVHDDKLEALKIDSIRIEBV-RPD-EVTLNDLARFTL 588
QY 215 KADVAQSSYGLYGCGQIYAVADTGLDTRNDSSM---HEAFRGKITLALYALGRNNAND 270
DB 589 NANTIALSTSYEGNGQKCVADTPGQKMADEMGILVHPANGVEHLEALM-LGDSKD 647
QY 271 TNGGTHVAGSVLNGKXTN-----KGMAPONLIVFQSI-----MDSKGLGLPSNL-QT 319
DB 648 TAGGTHYVCAICGGLYKNGDIRKRGVAPGATLMVQSIAGVSRDPNGKALEVPMDLGLQ 707
QY 320 LFSQAKSAGARIHTNSWGA---AVNG-AYTTSRNVDDY-RKNDMTILPAAGNEKXNG 373
DB 708 LFSNPYKLGRIHNSWGKWDATKAGQGLGEGQAMVDIDKFIYDHDFFVLVAAGNNAKA 767
QY 374 GT-----ISAPGTAKNAITVAGTENLRPSFGSYADN-----INHYAQSSTGPTD- 419
DB 768 KSKSNHIGAAQSAFNCITVAGTITRRPNNDYGPDNEVGAKPMTRINDTAKFSRGPYKPG 827

QY 420 -----GRKPDVMAFGTXILSARS-SIAPDS-----SFWANHDSKYAMGTSNATP 465
DB 828 RDINGNEYAGRIKPDVAVPAGVAILLSAASRAVAKQSRNRMVGRGDDMTWTSMTS 887
QY 466 IVAGNVQLRHFVNNRGITPKPSILKKAALJAGAAD-----XGIGPNGNGMGWVYLD 519
DB 888 LVAGCVMLLRBALKEHGEKRSALILKALVNGAVNPSQGLGLGTY-DYQGRGVRVID 946
QY 520 KSLNV---AVYNESSKISTSQ-----KATYKFTAT-----AGK-PLKISLWSD 559
DB 947 SSISWVKLSFVDDGKLFEDYQFVAPLRQVPEERERWTSLSLIVPAGRNRLTYLAVPD 1006
QY 560 APASTTASVTLPNDLIVTAPNGTXVYVNDPFXKPKXXNMGRNVNVPFINXQSSGYT 619
DB 1007 KPAQ---SGLMQNDINLVLVS-GAERHGNMKGP---GYDHTNVEKTIWENVPGETFK 1059
QY 620 IEVOAY-NVPY-GPOXFSIA 637
DB 1060 IVASIMWNIDVKAPTSFAVA 1079

RESULT 13
AC 054M84 D1CIDI
ID 054M84 D1CIDI PRELIMINARY; PRT; 1741 AA.
AC 054M84-
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE ABC transporter B family protein.
GN Name=tagC; ORFNames=DDB011192;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Suggang B., Bertram M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivoiro F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kexhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hanser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Lousaged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simonde M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winkler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kaye R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrett B., Kaya A.;
RT "The genome of the social amoeba Dictyostelium discoideum".
RL Nature 01-01(2005).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAF01000133; EAL64353.1; -; Genomic_DNA.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR000209; pep_t_S53.


```
FT COMBRIAS 1353 1357 Poly-Asn.
FT COMBRIAS 1358 1364 Poly-Asp.
FT COMBRIAS 1381 1386 Poly-Asn.
FT COMBRIAS 1707 1729 Poly-Asn.
FT ACT_SITE 325 325 Charge relay system (By similarity).
FT ACT_SITE 372 372 Charge relay system (By similarity).
FT ACT_SITE 637 637 Charge relay system (By similarity).
FT CARBOHYD 390 397 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 536 536 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 547 547 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 614 614 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 689 689 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 735 735 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 741 741 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 776 776 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 832 832 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 887 887 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1251 1251 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1385 1385 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1386 1386 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1454 1454 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1704 1704 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1743 AA; 194146 MW; 12DB363B2F729839 CRC64;

Query Match 16.9%; Score 518.5; DB 1; Length 1743;
Best Local Similarity 24.2%; Pred. No. 3e-28;
Matches 194; Conservative 114; Mismatches 213; Indels 281; Gaps 33;

QY 88 HIXQNGILRETKOXL-----EXTGAKIIDYIPDYAYIV---EYEGDVXK 131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 115 YIVGPKRINDETRQGLKEPLIGTDIVLDEQRYSHIYHYPHDSFLVLMQEOGVLLSS 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 132 XXXIEHVSEVPY-----LPXYIDPOLFTKGSXLYK--AKALDT--KQXN 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 175 KEWYWGIEFEPBNKHILNYNEKSIGLPYII-----LSDSTNSLIQWMENTLSIKSYN 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 175 KEVQLRGI-----EXIAQKXSNVYVYITKPEYK 204
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 231 SKVLTLLNQKKLSIYVYNDESPSPSCSLINSEKLYVQWISSEQSSN--FIERSEKQ 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 205 VMDNVARGIV-----KADVAQSSYGLYGQGIIVAAVDITLDGR--NDS----- 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 288 TANLSPRVVPGTQDTLVNNDVDP-----LRKGQGLISTADTGLDSHCFPSISKPI 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 247 ---SMHEAFRKITALVALGRNNANDNGHGTAVASVLG-----NGXTNKGAPQA 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 343 PLNSVNLNHR--KVLYVITTSDDSDKDKYDGLVPLDAGARVHCDMSGVSVSGYSSDPTAS 401
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 297 NLVQSIWDSXGLGLG--PENVLTLPFQAKSAGARHITNSMGA---AVNGAYTTDSRN 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 402 KIAF--FDLASGSSSLTPPSDLKQLYPLDAGARVHCDMSGVSVSGYSSDPTAS 458
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 351 VDDYVRKN--DMTILFAAGNEXPNCGTIS--APGTAKNIYVQATENTLR-----PSRFS 400
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 459 IDDLPLFTHPDIILRAAGN---NEGYLSLTLQSTAKVITVIGAHQTHENTLVTDGPNNIN 515
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 401 YADNI----- 405
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 516 YQSSVLTINQELICDFSRNYCYTTAAOCCLASNATTTGLASCCPTLLRKSVIDAANTQPLLY 575
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 406 --NHVAQPSRSGPTKDGRIKPDVWAPGXIISARSSLA-----PDGSFMANHDSKVYM 457
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 576 NENNI CSFSSKGFTHDGMKPALVAPGYITYSARNGNMTTQCCDGL--DPTVALA-I 633
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 458 GGTSMATPIVAGNVAQLREH-----FVKRNGITPKPSLLKALAIAGA----- 499
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 634 SGTSMATSPFAAATTLIRQYLVDGYTPGSIYESKLOPTSGSLKALMINNAQLNGTFQ 693
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 500 --ADXLGLAYPNGN-----QGWGRVTLDSKLVNVAAYNESS----- 531
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 694 LITSSSIYPSNQVFENFAGASLVQGMGAIRMSNMLAHVNNNNNSNNNKTSDGITKFDGI 753
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 532 -----XLTSTQKATYXT-----ATAGKPRK-----ISLWMSDAPASTT 565
   : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```
DB 754 GGLDLRLVKNQWKESLSTQNTSYCTIYPSSSSSNSGNIRRVATLVWTPPSTAG 813
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 566 ASVLVNDLDELVI-----TAPNGTYVGNDFXXXKXNMDGRNVENVFINXP 613
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 814 AKPLVNLNLDITMILYRNDNGSTIRYSNQGSSFLG-----LAPQDTLVNVEGIVHNPT 867
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 614 QSGTYTTEVQAINVNPVGPQKFS 635
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 868 EPMYTRFVAVGTNVMPGPNPFS 889
   : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
Q8T9W1_DICDI PRELIMINARY; PRT; 1825 AA.
ID Q8T9W1_DICDI PRELIMINARY; PRT; 1825 AA.
AC Q8T9W1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Serine protease/ABC transporter tagp (ABC transporter B family
DE protein).
GN Name=tagp; ORFNames=ADB0191427;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxId=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugarcang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Plicher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kethothou A., Nle X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Fathothou P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Murty D., Mourlet T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchtleser C.,
RA Wadrop A., Felder M., Thangavelu M., Johnson D., Knutsen A.,
RA Loulsged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Splegler S., Rivey A.,
RA Shaulsky G., Schleicher M., Woodward J., Minkler T., Tanaka Y.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrett B., Kupea A.,
RA "The genome of the social amoeba Dictyostelium discoideum."
RL Nature 0:0-0(2005).
DR EMBL; AF466309; AAL74253.1; -; Genomic DNA.
DR EMBL; AAR0100033; EAL64354.1; -; Genomic DNA.
DR HSP; P08716; IMT0.
DR Dictybase; DDB0191427; tagp.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0046262; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR000439; ABC_transp_like.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00664; ABC_membrane_1.
DR Pfam; PF00005; ABC_tran; 1.

```

Search completed: April 7, 2006, 17:42:41
Job time : 152.117 secs

DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR ProDom; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS50929; ABC_TM1P; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE; PS00138; SUBTILASE_SER; UNKNOWN_1.
KW Peptidase.
SQ SEQUENCE 1825 AA; 202642 MW; E28160BC78613A3B CRC64;

Query Match 16.1%; Score 493; DB 2; Length 1825;

Best Local Similarity 24.2%; Pred. No. 2,3e-26;

Matches 192; Conservative 105; Mismatches 229; Indels 268; Gaps 31;

```
OY 88 HIXQNGPILEET---KQXLEXTG-----AKIIDIYDPYAYV---EYEGDVXSK 131
DB 127 YIVQEKDHINDETRQKQFLINTDIVLDEQPYQSHIVNYIPHDSFLVLMNDEQSNLSS 186
OY 132 XXXIHHVSEVER-----YLPXYXIDPOLFTKGASXL-----VK 164
DB 187 KEVWSWIGEFEPSENKIHLYNNEKSGLPYIKLSDSTNSLIQWENTLNSILTSYNSKYK 246
OY 165 AXALPTKQKNKEVQLRGIEIXQXXS-----NDVYIYAKPEYKVMNDV 209
DB 247 LTLINQKCL-KSIYVCNDESSSQSCSLVSEKLYYQWISEQSESNTYIERSEKFGPANRL 305
OY 210 ARGIV-----KADVAGSYGLYGCQIIVAVADTGLDTGR---NDS-----SM 248
DB 306 SPKALFGTMDTLVNNDRIDIP-----LRGKGQILSIADTGLDGSHPFSDSNMPIPNSV 360
OY 249 HEAFRGKITLALYALGRTNNANDTNGHGTIVAGSVLGNCKTN-----KGMAPQANLVFQ 301
DB 361 NLNMRKVVYTGSL--HDNEVDYDGHGTVCSSAAGAPESSLAISFSGLATDAKIAVF 418
OY 302 SI-MDSXGGLGGLPENLQTLFQOAXSAGARIHTNSMGA---AVNGAYTTDSRNVDDYR 356
DB 419 DIASDPSNNEPPEPFDYQOLYQPLYNAGNRVHGDMSGLSIGYLSYSDDAGSIDDFLY 478
OY 357 KN-DWTILPAAGNEXPNGTISAPGTAKNAITVGATENLRPSF----- 398
DB 479 THPDRILLRAAGNNEQYSSLS-QATAGNVITVGAHQTHESYTTDALREYSNFEVAKST 537
OY 399 -----GSYAD-----NINHYAQ 410
DB 538 LNSLCQSPDKKCTYTTAQCCTEYGVKGLSGCCTSYIKNSYASIFSSQPELYNENNICS 597
OY 411 PSSRGPTKDGRIKPPDMAPGTXILSARSLA-----PDSSFWANHDSKYAYMG 459
DB 598 PSSKQPTHDRKLPDIVAFQYITSAKSNAGATTDOCGDGLPNTNALISR-----SG 650
OY 460 TSMATPIVAGNVAGLREH-----FVKNRGITPKPSLLKALILAGADXLGYR-- 507
DB 651 TSMATPLATTAATTLIRQYLVGQYPTGSIYESNKLQPTGSLKALMINNAQLNNGTFPLS 710
OY 508 -----NGQGMGRVTLDKSLANAYNESS----- 531
DB 711 STNTNPSNAVDPDPAGANFVQGMGSLRMSFWL---YVSSSGYKPKPSRWVWGIGELGDKK 767
OY 532 -----XLSTSQKATYKFT-----ATAGKP-LKISLYWSDAPASTTASVTLVNDLD 575
DB 768 ASNKEYSLSLTSQNVSYCTTYKPPSSGSGNSGGIPRIVALVWTDPPSYSGAKLNLVNNID 827
OY 576 LVIT-----APNGTXVYG--NDPXXPXXNMNDGRNVENVP---INXQSGTYTIE 621
DB 828 LTMNTBESBFIIFYNSGSGSYNGTGTTLPLQ---DSINNVEGILYTPINTKSEISPRFI 884
OY 622 VQAYNVVGPQXFS 635
DB 885 IAGTNIPIGPQNFS 898
```

This Page Blank (uspto)


```
Db 241 TGRDSSMHEAFRGKITLALYALGRTNANDTNGHGTVAAGSVLGNSTNKGMAPOANLVF 300
Qy 301 QSIMDSXKGLGGLPSNLQTLFSGAASAGARLHTNSWGAAVNGAYTTDSRNVDDVYRKNDM 360
Db 301 QSIMDSXKGLGGLPSNLQTLFSGAASAGARLHTNSWGAAVNGAYTTDSRNVDDVYRKNDM 360
Qy 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHYAQSRSRPTXDG 420
Db 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHYAQSRSRPTXDG 420
Qy 421 RIKPDVMAPGTIIISARSSILAPDSSFWANHDSTKAYMGSTMATPIVAGNVAAQLREHFVK 480
Db 421 RIKPDVMAPGTIIISARSSILAPDSSFWANHDSTKAYMGSTMATPIVAGNVAAQLREHFVK 480
Qy 481 NRGITPKPSLLKAAIAGAADIGLGYPNGNGGWRVTLDKSLNVAAYNESSSLSTSOKAT 540
Db 481 NRGITPKPSLLKAAIAGAADIGLGYPNGNGGWRVTLDKSLNVAAYNESSSLSTSOKAT 540
Qy 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLDTTAPNGTXYGNDPXXKXKXND 600
Db 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLDTTAPNGTXYGNDPXXKXKXND 600
Qy 601 GRNNVENVFINKPOSGTTIEVOAYNVPGQPSLAIYN 640
Db 601 GRNNVENVFINKPOSGTTIEVOAYNVPGQPSLAIYN 640
```

RESULT 2

US-09-920-954-6

```
/ Sequence 6, Application US/09920954
/ Patent No. 6759228
/ GENERAL INFORMATION:
/ APPLICANT: TAKAIWA, MIKIO
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SAEKI, KATSUHISA
/ APPLICANT: KUBOTA, HIROMI
/ APPLICANT: HITOMI, JUN
/ APPLICANT: KAGEYAMA, YASUSHI
/ APPLICANT: SHIKATA, SHITSUW
/ APPLICANT: NOMURA, MASAFUMI
/ TITLE OF INVENTION: ALKALINE PROTEASE
/ FILE REFERENCE: 0327-0832-0PCT
/ CURRENT APPLICATION NUMBER: US/09/920,954
/ CURRENT FILING DATE: 2001-08-03
/ PRIOR APPLICATION NUMBER: 09/509,814
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: PCT/JP98/04528
/ PRIOR FILING DATE: 1998-10-07
/ PRIOR APPLICATION NUMBER: JP 9-274570
/ PRIOR FILING DATE: 1997-06-08
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 6
/ LENGTH: 640
/ TYPE: PRT
/ ORGANISM: Bacillus sp.
US-09-920-954-6
```

Query Match 99.0%; Score 3029; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 3e-258;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
Qy 1 MRXKKVFLSVLSAAAILSTVALKNPSAGAKXFDLDFKGIQTITDDXXGFSKXQGTGAAA 60
Db 1 MRXKKVFLSVLSAAAILSTVALKNPSAGAKXFDLDFKGIQTITDDXXGFSKXQGTGAAA 60
Qy 61 FLVSENVKLGKGLKLETPANNKLIHQFNPILEBTQXLEXTGAKILDIYIPDAY 120
Db 61 FLVSENVKLGKGLKLETPANNKLIHQFNPILEBTQXLEXTGAKILDIYIPDAY 120
Qy 121 IVEEGDVXKXKXLEHVESVEPYLPXYXIDPQLFTKGASLVKAKALDTKXNKXEVOLR 180
Db 121 IVEEGDVXKXKXLEHVESVEPYLPXYXIDPQLFTKGASLVKAKALDTKXNKXEVOLR 180
```

```
Db 121 IVEEGDVXKXKXLEHVESVEPYLPXYXIDPQLFTKGASLVKAKALDTKXNKXEVOLR 180
Qy 181 GIEKIAQXXXSNDYXYITAKEEYVMDVANGIYKADVAQSSYGLYGGQIIVAADTGLD 240
Db 181 GIEKIAQFAISNDVLYITAKEEYKVMDDVARGIVKADVAQSSYGLYGGQIIVAADTGLD 240
Qy 241 TGRDSSMHEAFRGKITLALYALGRTNANDTNGHGTVAAGSVLGNSTNKGMAPOANLVF 300
Db 241 TGRDSSMHEAFRGKITLALYALGRTNANDTNGHGTVAAGSVLGNSTNKGMAPOANLVF 300
Qy 301 QSIMDSXKGLGGLPSNLQTLFSGAASAGARLHTNSWGAAVNGAYTTDSRNVDDVYRKNDM 360
Db 301 QSIMDSXKGLGGLPSNLQTLFSGAASAGARLHTNSWGAAVNGAYTTDSRNVDDVYRKNDM 360
Qy 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHYAQSRSRPTXDG 420
Db 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHYAQSRSRPTXDG 420
Qy 421 RIKPDVMAPGTIIISARSSILAPDSSFWANHDSTKAYMGSTMATPIVAGNVAAQLREHFVK 480
Db 421 RIKPDVMAPGTIIISARSSILAPDSSFWANHDSTKAYMGSTMATPIVAGNVAAQLREHFVK 480
Qy 481 NRGITPKPSLLKAAIAGAADIGLGYPNGNGGWRVTLDKSLNVAAYNESSSLSTSOKAT 540
Db 481 NRGITPKPSLLKAAIAGAADIGLGYPNGNGGWRVTLDKSLNVAAYNESSSLSTSOKAT 540
Qy 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLDTTAPNGTXYGNDPXXKXKXND 600
Db 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLDTTAPNGTXYGNDPXXKXKXND 600
Qy 601 GRNNVENVFINKPOSGTTIEVOAYNVPGQPSLAIYN 640
Db 601 GRNNVENVFINKPOSGTTIEVOAYNVPGQPSLAIYN 640
```

RESULT 3

US-09-509-814A-8

```
/ Sequence 8, Application US/09509814A
/ Patent No. 6376227
/ GENERAL INFORMATION:
/ APPLICANT: TAKAIWA, MIKIO
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SAEKI, KATSUHISA
/ APPLICANT: KUBOTA, HIROMI
/ APPLICANT: HITOMI, JUN
/ APPLICANT: KAGEYAMA, YASUSHI
/ APPLICANT: SHIKATA, SHITSUW
/ APPLICANT: NOMURA, MASAFUMI
/ TITLE OF INVENTION: ALKALINE PROTEASE
/ FILE REFERENCE: 0327-0832-0PCT
/ CURRENT APPLICATION NUMBER: US/09/509,814A
/ CURRENT FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: PCT/JP98/04528
/ PRIOR FILING DATE: 1998-10-07
/ PRIOR APPLICATION NUMBER: JP 9-274570
/ PRIOR FILING DATE: 1997-06-08
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 8
/ LENGTH: 640
/ TYPE: PRT
/ ORGANISM: Bacillus sp.
US-09-509-814A-8
```

Query Match 99.0%; Score 3028; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 3.7e-258;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
Qy 1 MRXKKVFLSVLSAAAILSTVALKNPSAGAKXFDLDFKGIQTITDDXXGFSKXQGTGAAA 60
Db 1 MRXKKVFLSVLSAAAILSTVALKNPSAGAKXFDLDFKGIQTITDDXXGFSKXQGTGAAA 60
Qy 61 FLVSENVKLGKGLKLETPANNKLIHQFNPILEBTQXLEXTGAKILDIYIPDAY 120
Db 61 FLVSENVKLGKGLKLETPANNKLIHQFNPILEBTQXLEXTGAKILDIYIPDAY 120
```

```

Db      61 FLVSESNVLPKGLQKQLETFVPAANNKLIHQNGPILFETTKQLETKAKLIDYIPDVAY 120
Qy      121 IVEYEGDVYKXXXXIEHVESVEPYLPXYXIDPOLFTKGASXLVKAAALDTKXKNEVOQR 180
Db      121 IVEYEGDVYKASSTIEHVESVEPYLPYRIDPOLFTKGASELVKAAVALDTKXKNEVOQR 180
Qy      181 GIEIXAQXXXXSNDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGOIIVAVADTGLD 240
Db      181 GIEIXAQFPAISNDVLYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGOIIVAVADTGLD 240
Qy      241 TGRNDSMHBAFRGKITLALGRTNNANDTNGHGTVAAGSVLGNKXTKGMAPQANLYF 300
Db      241 TGRNDSMHBAFRGKITLALGRTNNANDTNGHGTVAAGSVLGNKXTKGMAPQANLYF 300
Qy      301 OSIMDSXGGLGGLPSNLQTLFSAQXASAGARIHTNSMGAAVNGAYTTDSHNDVDDYRKNDM 360
Db      301 OSIMDSXGGLGGLPSNLQTLFSAQXASAGARIHTNSMGAAVNGAYTTDSHNDVDDYRKNDM 360
Qy      361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
Db      361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
Qy      421 RIKPDVMAFGTYILSARSLAPDSSFMAHNSKYAVMGTSMATPIVAGNVAQLEHFPK 480
Db      421 RIKPDVMAFGTYILSARSLAPDSSFMAHNSKYAVMGTSMATPIVAGNVAQLEHFPK 480
Qy      481 NRGITPKPSILKAAIAGAADYGLGYPNGOGMGVTLDDKSLNVAVYNESSXLSTSQKAT 540
Db      481 NRGITPKPSILKAAIAGAADYGLGYPNGOGMGVTLDDKSLNVAVYNESSXLSTSQKAT 540
Qy      541 XYFTTAGPKPLKISLWMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXXPXXKXND 600
Db      541 XYFTTAGPKPLKISLWMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXXPXXKXND 600
Qy      601 GRNNVENVFINKPQSGTYTIEVOAYNVPGPQKFSIALYN 640
Db      601 GRNNVENVFINKPQSGTYTIEVOAYNVPGPQKFSIALYN 640

RESULT 4
US-09-920-954-8
/ Sequence 8, Application US/09920954
/ Patent No. 6759228
/ GENERAL INFORMATION:
/ APPLICANT: TAKAIWA, MIKIO
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SAEKI, KATSUHIISA
/ APPLICANT: KUBOTA, HIROMI
/ APPLICANT: HITOMI, JUN
/ APPLICANT: KAGEYAMA, YASUSHI
/ APPLICANT: SHIKATA, SHITSUM
/ APPLICANT: NOMURA, MASAFUMI
/ TITLE OF INVENTION: ALKALINE PROTEASE
/ FILE REFERENCE: 0327-0832-0PCT
/ CURRENT APPLICATION NUMBER: US/09/920, 954
/ PRIOR FILING DATE: 2001-08-03
/ PRIOR APPLICATION NUMBER: 09/509, 814
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: PCT/JP98/04528
/ PRIOR FILING DATE: 1998-10-07
/ PRIOR APPLICATION NUMBER: JP 9-274570
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 8
/ LENGTH: 640
/ TYPE: PRT
/ ORGANISM: Bacillus sp.
US-09-920-954-8

Query Match 99.0%; Score 3028; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 3.7e-258;
```

```

Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
Qy      1 MEXKKKKVFLSVSAAAIISTVALKNPISAGKAXFPILDFKGIOTTTDXKGFSSQXOTGAAA 60
Db      1 MEXKKKKVFLSVSAAAIISTVALKNPISAGKAXFPILDFKGIOTTTDXKGFSSQXOTGAAA 60
Qy      61 FLVSESNVLPKGLQKQLETFVPAANNKLIHQNGPILFETTKQLETKAKLIDYIPDVAY 120
Db      61 FLVSESNVLPKGLQKQLETFVPAANNKLIHQNGPILFETTKQLETKAKLIDYIPDVAY 120
Qy      121 IVEYEGDVYKXXXXIEHVESVEPYLPXYXIDPOLFTKGASXLVKAAALDTKXKNEVOQR 180
Db      121 IVEYEGDVYKASSTIEHVESVEPYLPYRIDPOLFTKGASELVKAAVALDTKXKNEVOQR 180
Qy      181 GIEIXAQXXXXSNDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGOIIVAVADTGLD 240
Db      181 GIEIXAQFPAISNDVLYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGOIIVAVADTGLD 240
Qy      241 TGRNDSMHBAFRGKITLALGRTNNANDTNGHGTVAAGSVLGNKXTKGMAPQANLYF 300
Db      241 TGRNDSMHBAFRGKITLALGRTNNANDTNGHGTVAAGSVLGNKXTKGMAPQANLYF 300
Qy      301 OSIMDSXGGLGGLPSNLQTLFSAQXASAGARIHTNSMGAAVNGAYTTDSHNDVDDYRKNDM 360
Db      301 OSIMDSXGGLGGLPSNLQTLFSAQXASAGARIHTNSMGAAVNGAYTTDSHNDVDDYRKNDM 360
Qy      361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
Db      361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
Qy      421 RIKPDVMAFGTYILSARSLAPDSSFMAHNSKYAVMGTSMATPIVAGNVAQLEHFPK 480
Db      421 RIKPDVMAFGTYILSARSLAPDSSFMAHNSKYAVMGTSMATPIVAGNVAQLEHFPK 480
Qy      481 NRGITPKPSILKAAIAGAADYGLGYPNGOGMGVTLDDKSLNVAVYNESSXLSTSQKAT 540
Db      481 NRGITPKPSILKAAIAGAADYGLGYPNGOGMGVTLDDKSLNVAVYNESSXLSTSQKAT 540
Qy      541 XYFTTAGPKPLKISLWMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXXPXXKXND 600
Db      541 XYFTTAGPKPLKISLWMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXXPXXKXND 600
Qy      601 GRNNVENVFINKPQSGTYTIEVOAYNVPGPQKFSIALYN 640
Db      601 GRNNVENVFINKPQSGTYTIEVOAYNVPGPQKFSIALYN 640

RESULT 5
US-09-509-814A-4
/ Sequence 4, Application US/09509814A
/ Patent No. 6376227
/ GENERAL INFORMATION:
/ APPLICANT: TAKAIWA, MIKIO
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SAEKI, KATSUHIISA
/ APPLICANT: KUBOTA, HIROMI
/ APPLICANT: HITOMI, JUN
/ APPLICANT: KAGEYAMA, YASUSHI
/ APPLICANT: SHIKATA, SHITSUM
/ APPLICANT: NOMURA, MASAFUMI
/ TITLE OF INVENTION: ALKALINE PROTEASE
/ FILE REFERENCE: 0327-0832-0PCT
/ CURRENT APPLICATION NUMBER: US/09/509, 814A
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: PCT/JP98/04528
/ PRIOR FILING DATE: 1998-10-07
/ PRIOR APPLICATION NUMBER: JP 9-274570
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 4
/ LENGTH: 639
/ TYPE: PRT
```

ORGANISM: Bacillus sp.
US-09-509-814A-4

Query Match 98.7%; Score 3020; DB 2; Length 639;
Best Local Similarity 93.4%; Pred. No. 1.9e-257;
Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
OY 4 KKKVFLSVLSAAALISTVALKNPSAGARXFDLPFKGIQTITDDXGFSKQOTGAAPLV 63
DB 3 KKKVFLSVLSAAALISTVALKNPSAGARXFDLPFKGIQTITDDXGFSKQOTGAAPLV 62
OY 64 ESENVKLLKGLKKLETPANNKLIHQFNGLILEETKQLEXTGAKILDIYIPYAYIVE 123
DB 63 ESENVKLLKGLKKLETPANNKLIHQFNGLILEETKQLEXTGAKILDIYIPYAYIVE 122
OY 124 YEGDVXSXXXIEHVESVEPYLPXYXIDPQLEFTKASXLVKXALDTKQNKVEVLRGIE 183
DB 123 YEGDVQSKVRSIEHVESVEPYLPXYXIDPQLEFTKASXLVKXALDTKQNKVEVLRGIE 182
OY 184 XIAQXXXSNDVXYITTAPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAADTGLDGR 243
DB 183 EIAQVASNDVXYITTAPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAADTGLDGR 242
OY 244 NDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVLGNKTNKGMAPOANLVFQSI 303
DB 243 NDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVLGNKTNKGMAPOANLVFQSI 302
OY 304 MDSKGLGGLPSNLQTLFSGQXSAGARIHTNSWGAAVNGAYTTDSRVDDYVRKNDMTIL 363
DB 303 MDSKGLGGLPSNLQTLFSGQXSAGARIHTNSWGAAVNGAYTTDSRVDDYVRKNDMTIL 362
OY 364 PAAGNEXPNGGTISAPGTAKAALITVGATENLRPSFGSYADININHAQFSSSGPKDGRIR 423
DB 363 PAAGNEXPNGGTISAPGTAKAALITVGATENLRPSFGSYADININHAQFSSSGPKDGRIR 422
OY 424 PDVNAFGTYILSARSSSLAPDSSFMANHDSKAYVNGTSMATPIVAGNVAQLREHFVKNRG 483
DB 423 PDVNAFGTYILSARSSSLAPDSSFMANHDSKAYVNGTSMATPIVAGNVAQLREHFVKNRG 482
OY 484 ITPKPSLLKALILAGAADXGIGYPNGNGQWGRVTLDKSLNVAAYNNESSSLSTSOKATYTF 543
DB 483 ITPKPSLLKALILAGAADXGIGYPNGNGQWGRVTLDKSLNVAAYNNESSSLSTSOKATYTF 542
OY 544 TATGKPKLISLWSDAPASTTASVTLVNDLDTVITAPNGTYVGNDFXXPXXNWDGRN 603
DB 543 TATGKPKLISLWSDAPASTTASVTLVNDLDTVITAPNGTYVGNDFXXPXXNWDGRN 602
OY 604 NVENVFINKPQSGTYTIEVOAYNVPGPQKFSLAIYN 640
DB 603 NVENVFINKPQSGTYTIEVOAYNVPGPQKFSLAIYN 639
```

RESULT 6
US-09-920-954-4
; Sequence 4, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-920-954-4

Query Match 98.7%; Score 3020; DB 2; Length 639;
Best Local Similarity 93.4%; Pred. No. 1.9e-257;
Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
OY 4 KKKVFLSVLSAAALISTVALKNPSAGARXFDLPFKGIQTITDDXGFSKQOTGAAPLV 63
DB 3 KKKVFLSVLSAAALISTVALKNPSAGARXFDLPFKGIQTITDDXGFSKQOTGAAPLV 62
OY 64 ESENVKLLKGLKKLETPANNKLIHQFNGLILEETKQLEXTGAKILDIYIPYAYIVE 123
DB 63 ESENVKLLKGLKKLETPANNKLIHQFNGLILEETKQLEXTGAKILDIYIPYAYIVE 122
OY 124 YEGDVXSXXXIEHVESVEPYLPXYXIDPQLEFTKASXLVKXALDTKQNKVEVLRGIE 183
DB 123 YEGDVQSKVRSIEHVESVEPYLPXYXIDPQLEFTKASXLVKXALDTKQNKVEVLRGIE 182
OY 184 XIAQXXXSNDVXYITTAPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAADTGLDGR 243
DB 183 EIAQVASNDVXYITTAPEYKVMNDVARGIVKADVAOSSYGLYGQGIIVAADTGLDGR 242
OY 244 NDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVLGNKTNKGMAPOANLVFQSI 303
DB 243 NDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVLGNKTNKGMAPOANLVFQSI 302
OY 304 MDSKGLGGLPSNLQTLFSGQXSAGARIHTNSWGAAVNGAYTTDSRVDDYVRKNDMTIL 363
DB 303 MDSKGLGGLPSNLQTLFSGQXSAGARIHTNSWGAAVNGAYTTDSRVDDYVRKNDMTIL 362
OY 364 PAAGNEXPNGGTISAPGTAKAALITVGATENLRPSFGSYADININHAQFSSSGPKDGRIR 423
DB 363 PAAGNEXPNGGTISAPGTAKAALITVGATENLRPSFGSYADININHAQFSSSGPKDGRIR 422
OY 424 PDVNAFGTYILSARSSSLAPDSSFMANHDSKAYVNGTSMATPIVAGNVAQLREHFVKNRG 483
DB 423 PDVNAFGTYILSARSSSLAPDSSFMANHDSKAYVNGTSMATPIVAGNVAQLREHFVKNRG 482
OY 484 ITPKPSLLKALILAGAADXGIGYPNGNGQWGRVTLDKSLNVAAYNNESSSLSTSOKATYTF 543
DB 483 ITPKPSLLKALILAGAADXGIGYPNGNGQWGRVTLDKSLNVAAYNNESSSLSTSOKATYTF 542
OY 544 TATGKPKLISLWSDAPASTTASVTLVNDLDTVITAPNGTYVGNDFXXPXXNWDGRN 603
DB 543 TATGKPKLISLWSDAPASTTASVTLVNDLDTVITAPNGTYVGNDFXXPXXNWDGRN 602
OY 604 NVENVFINKPQSGTYTIEVOAYNVPGPQKFSLAIYN 640
DB 603 NVENVFINKPQSGTYTIEVOAYNVPGPQKFSLAIYN 639
```

RESULT 7
US-09-509-814A-2
; Sequence 2, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE

```
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRF
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
```

```
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-2
```

Query Match

98.6%, Score 3016, DB 2, Length 640;

Best Local Similarity 100.0%; Pred. No. 4.2e-257;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRKKKKVFLSVLSAAALISTVALXNPSSAGAXPFLDPFGIGTDTTDXGFGSKOXGTGAAA 60
DB 1 MRKKKKVFLSVLSAAALISTVALXNPSSAGAXPFLDPFGIGTDTTDXGFGSKOXGTGAAA 60
QY 61 FLVESENVTLAKGLKKLETTVPANNKLAHXQPNGLLEETKQKLEXTGAKIIDYIPDAY 120
DB 61 FLVESENVTLAKGLKKLETTVPANNKLAHXQPNGLLEETKQKLEXTGAKIIDYIPDAY 120
QY 121 IVEYEGDVAXXXXXIEHVESVEPYL.PXXYIDQLFTKGASLVKXAXALDTKQXNEVQLR 180
DB 121 IVEYEGDVAXXXXXIEHVESVEPYL.PXXYIDQLFTKGASLVKXAXALDTKQXNEVQLR 180
QY 181 GLEXIAQXXXXNDVYITAKPEYKVMNDVARGIVRADVAQSSYGLYGQQLVAVADTGLD 240
DB 181 GLEXIAQXXXXNDVYITAKPEYKVMNDVARGIVRADVAQSSYGLYGQQLVAVADTGLD 240
QY 241 TGRNDSMHEAFRGKITLALYALGRNNANDTNGHGTAVAGSYLGNGXTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRNNANDTNGHGTAVAGSYLGNGXTNKGMAPOANLVF 300
QY 301 OSIMDSXGGLGGLPSNLQTLFSGOAXSAGARIHTNSGAAVNGAYTTDSRNDVYRKNDM 360
DB 301 OSIMDSXGGLGGLPSNLQTLFSGOAXSAGARIHTNSGAAVNGAYTTDSRNDVYRKNDM 360
QY 361 TLFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSVADNINHAQSSRGPYKDG 420
DB 361 TLFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSVADNINHAQSSRGPYKDG 420
QY 421 RIKPVMARGTITLARSLLAPDSSFANHDSKRYVMGTSNATPIVAGNVQLEHHPK 480
DB 421 RIKPVMARGTITLARSLLAPDSSFANHDSKRYVMGTSNATPIVAGNVQLEHHPK 480
QY 481 NRGITPKPSILKAALIAGAADKGLGYPNGNOCMGRTLDKSLNVYVNESSLSTSOXAT 540
DB 481 NRGITPKPSILKAALIAGAADKGLGYPNGNOCMGRTLDKSLNVYVNESSLSTSOXAT 540
QY 541 YXFTATAGPKLISLVMSDAPASTJASTVTLVNDLVLITAPNGTYVGNDFXXPXXKWD 600
DB 541 YXFTATAGPKLISLVMSDAPASTJASTVTLVNDLVLITAPNGTYVGNDFXXPXXKWD 600
QY 601 GRNNVENPINKPQSGTITIEVOATNVPPGPOXFSLATV 640
DB 601 GRNNVENPINKPQSGTITIEVOATNVPPGPOXFSLATV 640

RESULT 8
US-09-920-954-2
Sequence 2, Application US/09920954
Patent No. 6759228
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHI SA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/930,954
PRIOR APPLICATION NUMBER: 2001-08-03
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/J998/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0

SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (184)..(184)

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-920-954-2

Query Match 98.6%; Score 3016; DB 2; Length 640;
Best Local Similarity 100.0%; Pred. No. 4.2e-257;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEXKKVPLSTVSAALISTVALNPNPAGAKXPDLPFGIGTITDDXKPSKXQTGA 60
DB 1 MEXKKVPLSTVSAALISTVALNPNPAGAKXPDLPFGIGTITDDXKPSKXQTGA 60
QY 61 FLVSEENVTLKKKGLKKTETVPANNKGLHXQFNGLIESTKQXLEXTGAKILDIYIPDAY 120

DB 61 FLVSEENVTLKKKGLKKTETVPANNKGLHXQFNGLIESTKQXLEXTGAKILDIYIPDAY 120
QY 121 IVEYEGDVXSKXXXIHHVSVBPIYPXYXIDQLPFGKASXLYKXALDTKQNKVQJR 180
DB 121 IVEYEGDVXSKXXXIHHVSVBPIYPXYXIDQLPFGKASXLYKXALDTKQNKVQJR 180
QY 181 GLEXIQXXXNDVXYITAKPEKWNDDVARGIVKADVQSSYGLVGQGIYVAVDTGLD 240
DB 181 GLEXIQXXXNDVXYITAKPEKWNDDVARGIVKADVQSSYGLVGQGIYVAVDTGLD 240
QY 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGVAGSVLGNKXTKGNAPQANTVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGVAGSVLGNKXTKGNAPQANTVF 300
QY 301 GSIMDSXGGLGGLPSNLQTLFQOAKSAGARHTNSKGAAVNGAYTTDSNRVDDYRKNDY 360
DB 301 GSIMDSXGGLGGLPSNLQTLFQOAKSAGARHTNSKGAAVNGAYTTDSNRVDDYRKNDY 360
QY 361 TILFAAGNEXPNGGTISAPGTAKNAITVGATENTLPSFGSYADNINHVAQFSSRGPTKDG 420
DB 361 TILFAAGNEXPNGGTISAPGTAKNAITVGATENTLPSFGSYADNINHVAQFSSRGPTKDG 420
QY 421 RIKPDVMAFGTYILSARSSLAPDSSFWANHDSKYAVMGTSNATPIVAGNVAQLRHFVK 480
DB 421 RIKPDVMAFGTYILSARSSLAPDSSFWANHDSKYAVMGTSNATPIVAGNVAQLRHFVK 480
QY 481 NRGITPKESGLAALLAGAADYGLGYPNGNQGMRVTLDKSLNAVYNESSXLTSSQKAT 540
DB 481 NRGITPKESGLAALLAGAADYGLGYPNGNQGMRVTLDKSLNAVYNESSXLTSSQKAT 540
QY 541 YKFTTAGKPLKISLVWSPAPASTASVTLVNDLVLVTPAAGTYVGVGDDFXKPKXND 600
DB 541 YKFTTAGKPLKISLVWSPAPASTASVTLVNDLVLVTPAAGTYVGVGDDFXKPKXND 600
QY 601 GRNNVENVFINKPQSGTYTIEVOAYNVVPVGPQXFSIATVN 640
DB 601 GRNNVENVFINKPQSGTYTIEVOAYNVVPVGPQXFSIATVN 640

RESULT 9
US-09-509-814A-1
Sequence 1, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIRATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0337-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURES:
NAME/KEY: misc feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid

NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature

LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match 98.3%; Score 3007; DB 2; Length 639;

Beet Local Similarity 100.0%; Pred. No. 2.6e-256; Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KKKVFLSVLSAAALISTVALNPSGAXRXPDLDFKGIQTITDDXXGSPKXQGTGAAPLV 63
DB 3 KKKVFLSVLSAAALISTVALNPSGAXRXPDLDFKGIQTITDDXXGSPKXQGTGAAPLV 62
QY 64 ESENVKLGKGLKCLLETTPANNKCHIXQFNPIIETQXLEXGAKLIDYIPYATYVE 123
DB 63 ESENVKLGKGLKCLLETTPANNKCHIXQFNPIIETQXLEXGAKLIDYIPYATYVE 122
QY 124 YEGDVYKXXXXIIEHVESVEPLPYXXIDPOLFTGASGLVRAXXLDTPKXKKEVQLRGIR 183
DB 123 YEGDVYKXXXXIIEHVESVEPLPYXXIDPOLFTGASGLVRAXXLDTPKXKKEVQLRGIR 182
QY 184 XIAQXXKSNVXYITAKPEYRVMDVARGIVKADVAOSSYGLYGGGQIVAAVADTGLDTGR 243
DB 183 XIAQXXKSNVXYITAKPEYRVMDVARGIVKADVAOSSYGLYGGGQIVAAVADTGLDTGR 242
QY 244 NDSWHEAFRQKITRALVLAGTNNANDTNGGTHVAGSVLNGXTNKGMAQNALVFGSI 303
DB 243 NDSWHEAFRQKITRALVLAGTNNANDTNGGTHVAGSVLNGXTNKGMAQNALVFGSI 302

SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-1

Query Match 70.4%; Score 2155; DB 2; Length 434;
Best Local Similarity 96.3%; Pred. No. 1.7e-181;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 207 NDVARGIVKADVAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITATYALGRTN 266
DB 1 NDVARGIVKADVAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITATYALGRTN 60
QY 267 NANTNGHGTIVAGSVLGNKTKMGMAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSAQXS 326
DB 61 NANTNGHGTIVAGSVLGNKTKMGMAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSAQXS 120
QY 327 AGARHTNSWGAANVGAATTTDSRVDDYVRKNDMTILPAAGNEXPNGGTISAPGTAKNAI 386
DB 121 AGARHTNSWGAANVGAATTTDSRVDDYVRKNDMTILPAAGNEXPNGGTISAPGTAKNAI 180
QY 387 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIRPDVMAPGTXILSARSLAPDSF 446
DB 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIRPDVMAPGTXILSARSLAPDSF 240
QY 447 WANHSKYAYMGTSMTPIVAGNVAQLREHFVKRGTTPKPSILKALLAGADVGLGY 506
DB 241 WANHSKYAYMGTSMTPIVAGNVAQLREHFVKRGTTPKPSILKALLAGADVGLGY 300
QY 507 PNGNGMGRVTLDSKLVNAAVYNESSXLSQKATYXFTATGKPLKISLWSDAPASTTA 566
DB 301 PNGNGMGRVTLDSKLVNAAVYNESSXLSQKATYXFTATGKPLKISLWSDAPASTTA 360
QY 567 SVTLVNDLDVITAPNGTYVYVNDPFXKPKXKXNDGRNVENVFINKPQSGTYTIEVQAYN 626
DB 361 SVTLVNDLDVITAPNGTYVYVNDPFXKPKXKXNDGRNVENVFINKPQSGTYTIEVQAYN 420
QY 627 VPVGQFQFSLAIYN 640
DB 421 VPVGQFQFSLAIYN 434

RESULT 14
US-09-985-689A-2
Sequence 2, Application US/09985689A
Patent No. 6803222
GENERAL INFORMATION:

APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGETAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483US0
CURRENT APPLICATION NUMBER: US/09/985,689A
PRIORITY FILING DATE: 2002-07-01
PRIORITY APPLICATION NUMBER: JP P2000-355166
PRIORITY FILING DATE: 2000-11-22
PRIORITY APPLICATION NUMBER: JP P2001-114048
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin version 3.1
SEQ ID NO 2
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-2

Query Match 70.4%; Score 2155; DB 2; Length 434;
Best Local Similarity 96.3%; Pred. No. 1.7e-181;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 207 NDVARGIVKADVAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITATYALGRTN 266
DB 1 NDVARGIVKADVAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITATYALGRTN 60
QY 267 NANTNGHGTIVAGSVLGNKTKMGMAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSAQXS 326
DB 61 NANTNGHGTIVAGSVLGNKTKMGMAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSAQXS 120
QY 327 AGARHTNSWGAANVGAATTTDSRVDDYVRKNDMTILPAAGNEXPNGGTISAPGTAKNAI 386
DB 121 AGARHTNSWGAANVGAATTTDSRVDDYVRKNDMTILPAAGNEXPNGGTISAPGTAKNAI 180
QY 387 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIRPDVMAPGTXILSARSLAPDSF 446
DB 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIRPDVMAPGTXILSARSLAPDSF 240
QY 447 WANHSKYAYMGTSMTPIVAGNVAQLREHFVKRGTTPKPSILKALLAGADVGLGY 506
DB 241 WANHSKYAYMGTSMTPIVAGNVAQLREHFVKRGTTPKPSILKALLAGADVGLGY 300
QY 507 PNGNGMGRVTLDSKLVNAAVYNESSXLSQKATYXFTATGKPLKISLWSDAPASTTA 566
DB 301 PNGNGMGRVTLDSKLVNAAVYNESSXLSQKATYXFTATGKPLKISLWSDAPASTTA 360
QY 567 SVTLVNDLDVITAPNGTYVYVNDPFXKPKXKXNDGRNVENVFINKPQSGTYTIEVQAYN 626
DB 361 SVTLVNDLDVITAPNGTYVYVNDPFXKPKXKXNDGRNVENVFINKPQSGTYTIEVQAYN 420
QY 627 VPVGQFQFSLAIYN 640
DB 421 VPVGQFQFSLAIYN 434

RESULT 15
US-09-985-689A-6
Sequence 6, Application US/09985689A
Patent No. 6803222
GENERAL INFORMATION:

APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGETAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483US0
CURRENT APPLICATION NUMBER: US/09/985,689A
PRIORITY FILING DATE: 2002-07-01
PRIORITY APPLICATION NUMBER: JP P2000-355166
PRIORITY FILING DATE: 2000-11-22
PRIORITY APPLICATION NUMBER: JP P2001-114048
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin version 3.1
SEQ ID NO 6
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-6

Query Match 68.1%; Score 2082; DB 2; Length 434;
Best Local Similarity 91.5%; Pred. No. 4.6e-175;
Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 207 NDVARGIVKADVAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITATYALGRTN 266
DB 1 NDVARGIVKADVAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITATYALGRTN 60

```

Qy 267 NNDPTNGHGTVAAGSVLGNXGXTNKGMAPOANLVFOSIMDSXGIGGLPSNLQTLPSQAXS 326
Db 61 NNDPTNGHGTVAAGSVLGNXGXTNKGMAPOANLVFOSIMDSXGIGGLPSNVSTLPSQAYS 120
Qy 327 AGARLHTNSWGA VNGAYTTDSRNVDYVRKNDMTLLFAAGNEXPNGGTISAPGTAKNAI 386
Db 121 AGARLHTNSWGA PVNGAYTTDSRNVDYVRKNDMAVLFAAGNEXPNGGTISAPGTAKNAI 180
Qy 387 TVGATENTLRFPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAPGTYILSARSSLAPDSSF 446
Db 181 TVGATENTLRFPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAPGTYILSARSSLAPDSSF 240
Qy 447 WANHD$KXAYMGATSMATPIVAGNVAQLEHFPVKRGTTPKPSLLKALILAGADKGLGY 506
Db 241 WANHD$KXAYMGATSMATPIVAGNVAQLEHFPVKRGTTPKPSLLKALILAGATDGLGY 300
Qy 507 PNGNOGMRVTLDKSLNVA YVNESSXLST$OKATYXFTATAGKPLKISLWSDAPASTTA 566
Db 301 PNGNOGMRVTLDKSLNVA FVNETSLSL$TNOKATYSFTAQSGKPLKISLWSDAPASTSA 360
Qy 567 SYTLVNDLDELVTAPNGIXYVGNDFKXPKXXNMGGRNNYENVPINXPOS GTTIEVQAYN 626
Db 361 SYTLVNDLDELVTAPNGIXYVGNDFTA PYDNNWMDGRNNYENVPINAPQSGT YTVVEQAYN 420
Qy 627 VPVGPOXPSLAIYN 640
Db 421 VPQGPAPPSLAIYN 434

```

Search completed: April 7, 2006, 17:44:49
 Job time : 35.0266 secs

Y:

This Page Blank (uspto)

Query Match	9.04;	Score 276.5;	DB 6;	Length 802;
Best Local Similarity	21.84;	Fred. No. 8.6e-14;		
Matches 160;	Conservative	82;	Mismatches 226;	Indels 265; Gaps 31
QY	127	DVKSXXXXIIEHVESVEPYLPXXYIDPOLFTTGAGSLVAKALDT---	KONKKEVOLRGIS	183
Db	49	DVTSDDKQTTIVIVELKEKSLAEAKADGEKQTFQSLKTARSKALKTLKKAKVNRVD-RVFS		107
QY	184	XIAOXXXXNDVXYITTAPEYKVM-----	NDVAKCIYKAD	217
Db	108	GFPMKLPASELPKLLAYKEVAVVYPNATYKPDYVKGDVTLAADAIFYQMKRSAPFIAD		167
QY	218	VAOSSYGLGGQGLVAVADTGLDTRGD-SSMHEAPRGK--ITALVYALGRTNANDTG-		273
Db	168	QAMKS-GYTGKGIYVAIYDTCVDYTHEDLKNKFPYKGYDEVDNDYDFQSTFPTGDPRGGA		226
QY	274	--HGTWAGSVLNGXGXTNKGMAPOANI VFOSIMDSXGGLGGLPSNLQTLFSGQASAGARI		331
Db	227	TDHGTHTVAGFTIAANGQI-KGVAFPEATLLAYRLVLPGG--SGTETENVIAGIEKAAVADGAKV		283
QY	332	HTNSWGAUNGA--TTTDSRVNDVYVRKDDMTILFAANGXNGGTISARPTAKAALITGCA		390
Db	284	MNTSLGNSLNSPDYATSI--ALDDAMAGVAVTNSNGSPENNTVSGPSTRVLAISVGA		341
QY	391	TE-----		400

Db 342 SOLPYNEYSVTLPSYSKAWMGYOEKDLLEALNGEVELVEAGLQADDPSGDKVKGVA 401
Qy 401 -----YADINH----- 407
Db 402 VIORGVIFPVDKAKNAKAGAI GAVITNNATGEIEANWGMVAFPVTKLSKEEGELVQOI 461
Qy 408 -----VAOFSSRGPTKD-GRIKPDVMAFGTXILSARSLAPDSSF 446
Db 462 KEGKHSVVFSLDKGGLGETTILASFSSRGVMDTMWIKPDVSAFGVNIYSTIPT----- 514
Qy 447 WANHDSK---YAWMGTSMATPIYAGNVAQLREHFVKNRGITP--KPSLKAALLIAGAA 500
Db 515 ---HDPKRPYVGSGKQGTSMASPHYAGTAALIKQ-----AKPDMTEQIKGVLMNTAE 564
Qy 501 ---DXGLGYPNCGMGWRVTLDKSLNV-----AVYNSS 531
Db 565 KLTBNGKPLPNTQAGSIRIMELKASSIVTPGSHSYGTFLKDKGKQTKQAATTIN- 623
Qy 532 XLSTSQKA---TYXFTATAGKPLKISLWSD---APASTT---ASVTLVMDLIVITAPN 582
Db 624 -LSSHAKAYOLEYSPKGTG---ITVSGTERVVVPANQGTAAAKVTVNSAKTKAGTYE 677
Qy 583 GTXYGVNDPKXPKXXNMDSRNNV---NVFINXP-----QSGTYYTIEVQAY 625
Db 678 GTVYIRE-----DGRKVAERIPLLIYKEPDYPRVTSVTEBPAGAKQAGYITE--AY 725
Qy 626 NVVPGPOKFXSLAI 638
Db 726 -LPGAEELAPLV 737

RESULT 2

US-10-510-386-28
; Sequence 28, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Raemussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28
; LENGTH: 874
; TYPE: PRF
; ORGANISM: Bacillus licheniformis
US-10-510-386-28

Query Match 9.0%; Score 274.5; DB 6; Length 874;
Best Local Similarity 23.7%; Pred. No. 1.4e-13;
Matches 141; Conservative 85; Mismatches 253; Indels 117; Gaps 25;
Qy 14 AAATLSTVALKNPAGAKRFDLPKGIQTITDXXGFSKQGTGAAP-----LV 63
Db 213 AAPFISAKMLFNKSRDSDVYKDL---VOLKPLPADVAKNGQVSAYSTINDOKAISRLY 268
Qy 64 ESENVKLLKGLXKKLETYPANNKLHIXQFNGLPILEETKQLEKXG-AKILDYIPDYAYIV 122
Db 269 ETAAASVPEPLKQDQVAKD--IGIEQLTG---SKSAVLEKAGMATASSAPENRITV 323
Qy 123 EYE-----GDVYKXXXXIEHVESVPE-----YLPYXXID-POLFTGASXLVKAALD 169
Db 324 KLKGGKPGSPKSKAQ-SGVQALEPLGSKTAFKDMYVEMKESRSRGFRAAKQYQAA 382
Qy 170 TKQNKKEVQLRGIEIXIQA-XXXSNDVYXITAKPEYKMNVDVARGIVKADVAQSSGLGYG 228
Db 383 ASKIAKPEVEFVEQVOYEALSRDTQY---PYQWLSKNGKNRANAANDIOFEOLOKLMK 439

Qy 229 GQ-----IYAVADTGLDTRNDSSMHEAFRGKI---TALVALGRTNNAINDTNGHTVAG 280
Db 440 GKGLKQTVIAVVDIGVHTLADLS-----GSVKKDEGVNVAQTADMDNGHTVAG 493
Qy 281 SVLG---NGXTNKGMARQANI VFOSSINDSKGLGGLBSNIOQLFSAQXSAGARIHTNSWG 337
Db 494 IIAAODHFMAGINNAVAKILPVKVLDSG--SGDTEQIANGIITYADHGAKYINISLG 551
Qy 338 AAVNGAYTDBSRNDDVYKRNKMTILPFAAGEXHNGGITSAPGAKNAITGATENLRPS 397
Db 552 ---GPSYRVEYALKTAASKNVITVAATGNDVGS--EISTPASSKTYLSGATNNL--- 602
Qy 398 FGSYADININHAQFSSRGPTKDGRIKPDVMAFGTXILSARSLAPDSSFMANHDSKYAAM 457
Db 603 -----DLVSDYSNKGKL-----DMVAPGTDI-----PSLVDPDN-----VTYM 636
Qy 458 GGTSMATPIYAGNVAQLREHFVKNRGITPSPSLKAAALLIAGADXLG---GYPNGN--- 510
Db 637 SGTSMALPHVAAAAGL-----LSQNPILKPKQIASILITETTDVADVAFEQDNPDPDYDLD 691
Qy 511 -----OGMGRVTLDKSLNVAYNVSSKLSQKATYXFTATACKPLKI 553
Db 692 IEPAAQIPGDFVSGMGLNVFHAASVPELNMKVHPVLRHTAVTGTAKSGVTYKI 747

RESULT 3

US-10-510-386-200
; Sequence 200, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Raemussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 200
; LENGTH: 1047
; TYPE: PRF
; ORGANISM: Bacillus licheniformis
US-10-510-386-200

Query Match 9.0%; Score 274.5; DB 6; Length 1047;
Best Local Similarity 23.7%; Pred. No. 1.8e-13;
Matches 141; Conservative 85; Mismatches 253; Indels 117; Gaps 25;
Qy 14 AAATLSTVALKNPAGAKRFDLPKGIQTITDXXGFSKQGTGAAP-----LV 63
Db 213 AAPFISAKMLFNKSRDSDVYKDL---VOLKPLPADVAKNGQVSAYSTINDOKAISRLY 268
Qy 64 ESENVKLLKGLXKKLETYPANNKLHIXQFNGLPILEETKQLEKXG-AKILDYIPDYAYIV 122
Db 269 ETAAASVPEPLKQDQVAKD--IGIEQLTG---SKSAVLEKAGMATASSAPENRITV 323
Qy 123 EYE-----GDVYKXXXXIEHVESVPE-----YLPYXXID-POLFTGASXLVKAALD 169
Db 324 KLKGGKPGSPKSKAQ-SGVQALEPLGSKTAFKDMYVEMKESRSRGFRAAKQYQAA 382
Qy 170 TKQNKKEVQLRGIEIXIQA-XXXSNDVYXITAKPEYKMNVDVARGIVKADVAQSSGLGYG 228
Db 383 ASKIAKPEVEFVEQVOYEALSRDTQY---PYQWLSKNGKNRANAANDIOFEOLOKLMK 439
Qy 229 GQ-----IYAVADTGLDTRNDSSMHEAFRGKI---TALVALGRTNNAINDTNGHTVAG 280
Db 440 GKGLKQTVIAVVDIGVHTLADLS-----GSVKKDEGVNVAQTADMDNGHTVAG 493
Qy 281 SVLG---NGXTNKGMARQANI VFOSSINDSKGLGGLBSNIOQLFSAQXSAGARIHTNSWG 337

494 IIAAODNHFMSAGINAAKILPVKYLDSG--SGDTEQIANGIYAADHAKVINTLSIG 551
338 AAVNGAYTTDSNVDYVAKNDMTLLFAAGNEXPNGCTISAPGTAKNAITVGTATENTLRPS 397
552 ----GPTSVMEYALKTAASKVNTTIVAATGNDGVS--ELSTYPASSKYTLISVGAATNML--- 602
338 FGSYADNINHNVAQFSRGPCKDGRIKPDVMAFGTXILSARSSLPDSSFWMHDSKCYAM 457
603 ----DLVSDYDNYGKGL-----DNVAPGTDL---PSLVDPGN-----VYTM 636
458 GGTSMATPIVAGNVAQLEHFPVKNNGITPKPSILAKAALAGAADKGL---GYPNGN--- 510
637 SGTSMAPAHVAAGL-----LSQNPILKPKQIASILTETTADVAFEEQDNPNPDYDLD 691
511 -----OQMGRVTLDSKILNVAIVNESSKLSSTQKATYKFTANGRPLKI 553
692 IEPAAQIRGYDFVSGWGRNLNVAHSAVPELNMKVPVNLNHTAVTGAAGVTVKI 747

RESULT 4

US-11-156-062-23
; Sequence 23, Application US/11156062
; Publication No. US2005028173A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PEROXYDOLASE ACTIVITY
; FILE REFERENCE: HEIN-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-11-156-062-23

Query Match 7.9%; Score 242.5; DB 7; Length 379;
Best Local Similarity 23.4%; Pred. No. 1.5e-11;
Matches 116; Conservative 58; Mismatches 166; Indels 155; Gaps 18;

QY 1 MRKKKVLSTLSAAIISTVALKNPSAKX--ARKFDLPFGKIGITTTDXKGFSGKXQGTGA 58
DB 1 MRRKKSFWLGMILTALPMLVFTNAFSDSASAAQPAKKNVEKDY-----IVGFKSGVKTAS 52
QY 59 AAFVSESNVQLKXGLKLETVPANNNLHIXQFNGPILEETKOKLXETGAKILDIYIDY 118
DB 53 ----VKQIITRESGKVDK-----QFR--IINAARLIDREALKVEKNDDPV 93
QY 119 AYIYEEGDVAXXXXXIEHVESVEPYLPYXXIDPOLFTKGASXLVAXALDTRKXNKEVQ 178
DB 94 AYVER-----DHVAHALAQTVFYGIP----- 114
QY 179 LRGEIXIAKXXKXNDVXXITTAPEYKVNNDVARGIVKADVAGSSYGLVGQGIYAVADTG 238
DB 115 -----LIKADKVQAQ--GFKANVYKVAVLDTG 139
QY 239 LDTGRNDSMHEAFRKGITLALGRTNNANDTNGHGVAGSV--LNGKXTNKGMAPOA 296
DB 140 IQASHPELDLNVGCG-----ASFVAGEAYN--TDNGHGHVAGTVALADNTTGVLDGVAHSV 192
QY 297 NLVPOSINDXKGLGLPSNLQTLPSQAXSAGARLHTHSWGAANVAGAYTTDSRNVDYVR 356
DB 193 SLVAVKVLNNSG--SGSYSGIVSGIEWATTKNDVYNNSLGGA---SGSTAMKQAVNDVAY 247

QY 357 KNDMTILFAAGNEXPNG--GTISAPGTAKNAITVGTATENTLRPSFGSYADNINHNVAQFSR 414
DB 248 ARGVVVAALNAGSSGNTNTTIGYPAKIDSVTAAGA-----VDSNINRASFSSV 296
QY 415 GPTKDGRIKPDVMAFGTXILSARSSLPDSSFWMHDSKCYAMGTSNATPIVAGNVA-- 472
DB 297 G-----ALEVMAQAGVSYTPY-----HTYATLNGTSNAPVHAGAAALLI 338
QY 473 -----QLEHFPVKNR 482
DB 339 LSKHPLSAGVYRNR 353

RESULT 5

US-11-020-602-236
; Sequence 236, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Batell, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GCS27C2
; CURRENT APPLICATION NUMBER: US/11/020,602
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 236
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: Hybrid of
; OTHER INFORMATION: Bacillus lichenus and Bacillus amyloliquefaciens
US-11-020-602-236

Query Match 7.9%; Score 242; DB 7; Length 272;
Best Local Similarity 31.9%; Pred. No. 1e-11;
Matches 84; Conservative 30; Mismatches 101; Indels 48; Gaps 10;

QY 214 VKADVAGSSYGLYGGQIVAAVADTGLDGRNDSMHEAFRKGITLALGRTNNANDTNG 273
DB 11 VQAPAHNR--GLTSGGVAVAVLDGIST-----HPLNLRGASFPGBE--STQDNG 61
QY 274 HGTHVAGSV--LNGXTNKGAPQANLVFOSIMSXGLGLPSNLQTLPSQAXSAGARI 331
DB 62 HGTHVAGTIALNNSIGVLTGVAPELVAHVVLGASG--SGSVSSIAQGLEWAGNNGMHV 119
QY 332 HTNSGVAIVNAGAYTTDSNVDYVAKNDMTILFAAGNEXPNG--GTISAPGTAKNAITYG 389
DB 120 INMSLGGSSNAL-----FAAVDKAVASGVVVAALNAGNGTSGSSSTVIGPKYRPSVING 175
QY 390 ATENLRPSFGSYADNINHNVAQFSRGPCKDGRIKPDVMAFGTXILSARSSLPDSSFWAN 449
DB 176 A-----VDSNNGRASFSVGP-----ELDWAPVSVISQSTLRG----- 208
QY 450 HDSKYAMGTSNATPIVAGNVA 472
DB 209 --NKYGAYNGTSMASPHVAGAAA 229

RESULT 6

US-11-020-602-6
; Sequence 6, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Batell, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND

```
;; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
;; FILE REFERENCE: GCS27C2
;; CURRENT APPLICATION NUMBER: US/11/020,602
;; PRIOR FILING DATE: 2004-12-22
;; PRIOR APPLICATION NUMBER: US 09/500,135
;; PRIOR FILING DATE: 2000-02-08
;; PRIOR APPLICATION NUMBER: US 09/060,872
;; PRIOR FILING DATE: 1998-04-15
;; NUMBER OF SEQ ID NOS: 240
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO: 6
;; LENGTH: 269
;; TYPE: PRN
;; ORGANISM: Bacillus lentus
US-11-020-602-6

Query Match      7.8%; Score 239; DB 7; Length 269;
Best Local Similarity 30.7%; Pred. No. 1,8e-11;
Matches 90; Conservative 30; Mismatches 91; Indels 82; Gaps 14;

QY 214 VKADVAGSSYGLYGGQIVAAVADTGLDGRNDSMHEAFRGKITAYALGRTNANDTNG 273
DB 11 VQAPAAHNR-GLTSGGVKVAVLDTGIST-----HPDLNIRGASFPVGEF-STQDNG 61
QY 274 HGTIVAGSV--LGNGXTKMGAPQANLVFQSIMDSXGGLGSPSLQTLFQAMXAGARI 331
DB 62 HGTIVAGTIALNNSTIGLVGAPSAELYAKVLGASG--SGSVSSIAOGLWAGNNGMTV 119
QY 332 HTNSGCA-----AVGAYTTDSRVNDVVRKNDITLTPAAGNEXPNGSTISAPGTA 382
DB 120 ANLSIGSPSPATLEQANNSA---TSRGV-----LVVAASGNS--GAGSISYPARY 165
QY 383 KNAITVGATE--NLRPSTGSIADNINHVAGFSSRGPTGGRKIPDVAPGTIILSARSSL 440
DB 166 ANAAVAGATDQNNNAASFQYAGGL-----DIVAPGVVQSTYPG- 205
QY 441 APDSFMANHDSKAYVMGTSMATPIVAGNVA-----OLRHFVYN 481
DB 206 -----STYASLNGTSMATPHVAGAAALVYKQKPNMSNVQIRNH-LKN 246

RESULT 7
US-11-020-602-2  #
; Sequence 2, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Bectell, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GCS27C2
; CURRENT APPLICATION NUMBER: US/11/020,602
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 382
; TYPE: PRN
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (163)...(163)
; OTHER INFORMATION: Xaa = Asn or Pro
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (164)...(164)
; OTHER INFORMATION: Xaa = Pro or Asn
; FEATURE:
; NAME/KEY: VARIANT
```

```
;; LOCATION: (168)...(168)
;; OTHER INFORMATION: Xaa = Asn or Asp
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (195)...(195)
;; OTHER INFORMATION: Xaa = Ala or Ser
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (196)...(196)
;; OTHER INFORMATION: Xaa = Ser or Ala
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (205)...(205)
;; OTHER INFORMATION: Xaa = Ala or Asp
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (206)...(206)
;; OTHER INFORMATION: Xaa = Asp or Ala
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (265)...(265)
;; OTHER INFORMATION: Xaa = Thr or Ser
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (266)...(266)
;; OTHER INFORMATION: Xaa = Ser or Thr
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (358)...(358)
;; OTHER INFORMATION: Xaa = Gln or Glu
US-11-020-602-2

Query Match      7.5%; Score 229; DB 7; Length 382;
Best Local Similarity 23.6%; Pred. No. 1.7e-10;
Matches 117; Conservative 40; Mismatches 151; Indels 188; Gaps 17;

QY 5 KKVELSVLSAAAILSTVALXNPSAGXARXPLDPKGIQTTDXGFSKQXQTGAAPLVE 64
DB 4 KKVWISLFLPALALFTMAFGSTSAQAAG-----KSNQEKRYIVGFKOTMSTMSAA----- 54
QY 65 SENVKKXGKLKLETPANNKHLIXQFNPILETKOXLEXTAKILDIYIPDYAYIVEY 124
DB 55 -----KKQDVISEKGVQKQFK--YVDAASATLNEKAVKELKQPSVAYVEE- 100
QY 125 EGDVXSKXXKIEHY-----BSVEPYLPYXTIDPOLFTKGASXLVKKALADTKQAKKEVQL 179
DB 101 -----DHYAHVAGSV-PY-----GVSQ-IKAPAL----- 123
QY 180 RGIEKIAQXXXXSNVXYITAKPEYKMNDDVARGIVKADVAGSSYGLVGGQIVAAVDTGL 239
DB 124 -----HSGGYTGSNVKVAVIDSGI 142
QY 240 DTGRNDSMHEAFRGKITAYALGRTNANDTNGHGTIVAGSV--LGNGXTKMGAPQAN 297
DB 143 D-----SSHPLDAVAGGASVWPSETKXFPQXNHGHTVAGTAAIANNISIGLVAPSPXX 196
QY 298 LVFQSIMDSXGGLGGLPSNLQTLFQAMXAGARLHTMSGCAVNVGATYTSRANDV----- 352
DB 197 LVAAVKVLGXKG-----SGQYSWINGIEWALNNNDVINMS 232
QY 353 -----DYVRKNDMTILPAAGNEXPNG--GTISAPGTAKNAITVGATENLRP 396
DB 233 LGGSGSAAALKAANDKAVASGVVVVAAAGNGKXGSSSTVGPGRKYSVIVAG----- 286
QY 397 SFGSYADNINHVAGFSSRGPTKGRKIPDVAPGTIILSARSLAPDSFMANHDSKXAY 456
DB 287 -----VDSNQRAAPSSVGP-----ELDVAPGVISQISTLRG-----NKKGA 323
QY 457 MGTSMATPIVAGNVA 472
DB 324 YNGTSMASPHVAGAAA 339
```

RESULT 8
US-11-156-062-14
; Sequence 14, Application US/11156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Sueanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PEROXYDOLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-14

Query Match 7.4%; Score 227.5; DB 7; Length 274;
Best Local Similarity 30.2%; Pred. No. 1.4e-10;
Matches 85; Conservative 35; Mismatches 108; Indels 53; Gaps 11;
QY 213 IYKADVAQSSYGLYGGQIVAVADTGLDGRNDSMHEAFRGKITLALGRTNNANDTN 272
DB 10 LKADKVAQ--GPKANVAVAVLDITGICASHPLDNLVVG-----ASFVAGAYNA-DGN 61
QY 273 GHGTHVAGSV--LGNGXTKKGMAPOANLVFQSIMDSXGGLGSLPSNLQTLFSGAXSAGAR 330
DB 62 GHGTHVAGTVALDITGTVLGVAFSVSLYAVVYLVSSG--SSSYGIVSGIEMATTNGMD 119
QY 331 IHTNSGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPG--GTISAPGTAKNAITV 388
DB 120 VTNMSLGGA---SGSTAMKQAVDNVAVRGVVVAAAAGSSGSGNTTIGYPAKYDSVLAV 176
QY 389 GATERLDRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMAPTXILSARSSILAPSSFWA 448
DB 177 GA-----VDSNSRFAFSFSSVG-----AELEWAPAGAYSTPT----- 210
QY 449 NHDSKYAVMGTSMTPTIVAGNVA-----QLREHFVKNR 482
DB 211 ---NTYATWNGTSMASPHVAGAAALILSKHPNLSASQVNR 248

RESULT 9
US-11-065-943-54
; Sequence 54, Application US/11065943
; Publication No. US20050250131A1
; GENERAL INFORMATION:
; APPLICANT: JESTIN, JEAN-LUC
; APPLICANT: VICHIER-GUERRE, SOPHIE
; APPLICANT: FERRIS, STEPHANE
; TITLE OF INVENTION: METHODS FOR OBTAINING THERMOSTABLE ENZYMES, DNA POLYMERASE I
; TITLE OF INVENTION: VARIANTS FROM THERMUS AQUATICUS HAVING NEW CATALYTIC ACTIVITIES,
; FILE REFERENCE: 266426USOXCIIP
; CURRENT APPLICATION NUMBER: US/11/065,943
; CURRENT FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: US 10/787,219
; PRIOR FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 54
; LENGTH: 275

; TYPE: PRT
; ORGANISM: Bacillus amyloliquefaciens
US-11-065-943-54

Query Match 7.4%; Score 227.5; DB 7; Length 275;
Best Local Similarity 28.7%; Pred. No. 1.4e-10;
Matches 81; Conservative 26; Mismatches 92; Indels 83; Gaps 9;
QY 214 VKADVAQSSYGLYGGQIVAVADTGLDGRNDSMHEAFRGKITLALGRTNNANDTN 273
DB 11 IKAP-ALHSQGYTGNVAVAVDSID-----SSHPLDKVAGASWMPSEETNPQDNN 63
QY 274 HGTHVAGSV--LGNGXTKKGMAPOANLVFQSIMDSXGGLGSLPSNLQTLFSGAXSAGAR 331
DB 64 HGTHVAGTVALDITGTVLGVAFSVSLYAVVYLVGADG----- 100
QY 332 HTNSGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPN 372
DB 101 -SGQYSWIIINGIEMALANNMDVINMSLGPGSSAALKAADVAVASGVVVAAAAGNBT 159
QY 373 G--GTISAPGTAKNAITVGAATENLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMA 430
DB 160 GSSSTVGYPGKYPYIANGA-----VDSNSORASFSSVGP-----ELDVMAFG 202
QY 431 TXILSARSSILAPSSFWANHDSKYAVMGTSMTPTIVAGNVA 472
DB 203 VSIQSTLPG-----NKGAVNGTSMASPHVAGAAA 232

RESULT 10
US-11-020-602-3
; Sequence 3, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Beteil, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GC527C2
; CURRENT APPLICATION NUMBER: US/11/020,602
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Bacillus amyloliquefaciens
US-11-020-602-3

Query Match 7.4%; Score 227.5; DB 7; Length 275;
Best Local Similarity 28.7%; Pred. No. 1.4e-10;
Matches 81; Conservative 26; Mismatches 92; Indels 83; Gaps 9;
QY 214 VKADVAQSSYGLYGGQIVAVADTGLDGRNDSMHEAFRGKITLALGRTNNANDTN 273
DB 11 IKAP-ALHSQGYTGNVAVAVDSID-----SSHPLDKVAGASWMPSEETNPQDNN 63
QY 274 HGTHVAGSV--LGNGXTKKGMAPOANLVFQSIMDSXGGLGSLPSNLQTLFSGAXSAGAR 331
DB 64 HGTHVAGTVALDITGTVLGVAFSVSLYAVVYLVGADG----- 100
QY 332 HTNSGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPN 372
DB 101 -SGQYSWIIINGIEMALANNMDVINMSLGPGSSAALKAADVAVASGVVVAAAAGNBT 159
QY 373 G--GTISAPGTAKNAITVGAATENLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMA 430
DB 160 GSSSTVGYPGKYPYIANGA-----VDSNSORASFSSVGP-----ELDVMAFG 202

QY 431 TXILASBSLAPDSSFWANHDSKYAMGTSMTATPIVAGNVA 472
 Db 203 VSIOSTPLPG-----NKEYANGTSMASPHVAGAAA 232

RESULT 11 US-11-156-062-12

Sequence 12, Application US/11156062
 Publication No. US20050281773A1
 GENERAL INFORMATION:
 APPLICANT: Wieland, Susanne
 APPLICANT: Polanyi-Bald, Laura
 APPLICANT: Prueser, Inken
 APPLICANT: Stehr, Regina
 APPLICANT: Maurer, Karl-Heinz
 TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PEROXYDOLASE ACTIVITY
 FILE REFERENCE: HENK-0134 / H5698
 CURRENT APPLICATION NUMBER: US/11/156,062
 CURRENT FILING DATE: 2005-06-17
 PRIOR APPLICATION NUMBER: PCT/EP2003/014127
 PRIOR FILING DATE: 2003-12-20
 PRIOR APPLICATION NUMBER: DE 102 60 903.9
 PRIOR FILING DATE: 2002-12-20
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.3
 SEQ ID NO 12
 LENGTH: 274
 TYPE: PRT
 ORGANISM: Artificial
 FEATURE:
 OTHER INFORMATION: Synthetic Construct
 US-11-156-062-12

Query Match 7.4%; Score 226.5; DB 7; Length 274;
 Best Local Similarity 30.2%; Pred. No. 1.7e-10;
 Matches 85; Conservative 34; Mismatches 109; Indels 53; Gaps 11;

QY 213 IVKADVAOSSGYLGQGIYAVADTGLDTRGSDSMHEAFRGKITALYALGRTNNANDTN 272
 Db 10 LIRKADKVOAQ-GFGKANVKAVALDTGLOASHPDLNVGG-----ASFVAGEAYN-TDGN 61
 QY 273 GHGTHVAGSV--LNGXTNKGMAPQANLVFQSIIMDSXGGLGSLPSNLQTLFSSQXSGAR 330
 Db 62 GHGTHVAGTVAALNTTGTGVLGAPSVSLYAVKVNSSG--SGSYSGIVSGIEMATTNGMD 119
 QY 331 IHTNSWGAAVGAYTTDSRVDDYVRKNDMTILFAAGNEXPNG--GTISAPGTAKNAITV 388
 Db 120 VINNSLGA---SGSTAKQAVDNAYARGVAVVAAAGNSGSGSTNTTIGYPAKYDSVIAY 176
 QY 389 GATENLRPFSGSYADNINHVAFSSRGPTKDGRIKPDYMAFGTILSARSSLAPDSSFWA 448
 Db 177 GA-----VDSNSNRASFSSVG-----ALEVMAFGAGVSTYPT----- 210
 QY 449 NHDSKYAMGTSMTATPIVAGNVA-----QLREHFYKUR 482
 Db 211 ---NTYATNGTSMASPHVAGAAALLISKPNLSASOVRN 248

RESULT 12
 US-11-020-602-5
 Sequence 5, Application US/11020602
 Publication No. US20060024764A1
 GENERAL INFORMATION:
 APPLICANT: Estell, David
 APPLICANT: Harding, Fiona
 TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
 TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
 FILE REFERENCE: GCS27G2
 CURRENT APPLICATION NUMBER: US/11/020,602
 CURRENT FILING DATE: 2004-12-22
 PRIOR APPLICATION NUMBER: US 09/500,135
 PRIOR FILING DATE: 2000-02-08
 PRIOR APPLICATION NUMBER: US 09/060,872

;; PRIOR FILING DATE: 1998-04-15
 ;; NUMBER OF SEQ ID NOS: 240
 ;; SOFTWARE: PatentIn Ver. 2.1
 ;; SEQ ID NO 5
 ;; LENGTH: 274
 ;; TYPE: PRT
 ;; ORGANISM: Bacillus licheniformis
 US-11-020-602-5

Query Match 7.4%; Score 225.5; DB 7; Length 274;
 Best Local Similarity 29.9%; Pred. No. 2.1e-10;
 Matches 84; Conservative 36; Mismatches 108; Indels 53; Gaps 11;

QY 213 IVKADVAOSSGYLGQGIYAVADTGLDTRGSDSMHEAFRGKITALYALGRTNNANDTN 272
 Db 10 LIRKADKVOAQ-GFGKANVKAVALDTGLOASHPDLNVGG-----ASFVAGEAYN-TDGN 61
 QY 273 GHGTHVAGSV--LNGXTNKGMAPQANLVFQSIIMDSXGGLGSLPSNLQTLFSSQXSGAR 330
 Db 62 GHGTHVAGTVAALNTTGTGVLGAPSVSLYAVKVNSSG--SGSYSGIVSGIEMATTNGMD 119
 QY 331 IHTNSWGAAVGAYTTDSRVDDYVRKNDMTILFAAGNEXPNG--GTISAPGTAKNAITV 388
 Db 120 VINNSLGA---SGSTAKQAVDNAYARGVAVVAAAGNSGSGSTNTTIGYPAKYDSVIAY 176
 QY 389 GATENLRPFSGSYADNINHVAFSSRGPTKDGRIKPDYMAFGTILSARSSLAPDSSFWA 448
 Db 177 GA-----VDSNSNRASFSSVG-----ALEVMAFGAGVSTYPT----- 210
 QY 449 NHDSKYAMGTSMTATPIVAGNVA-----QLREHFYKUR 482
 Db 211 ---NTYATNGTSMASPHVAGAAALLISKPNLSASOVRN 248

RESULT 13
 US-11-156-062-4
 Sequence 4, Application US/11156062
 Publication No. US20050281773A1
 GENERAL INFORMATION:
 APPLICANT: Wieland, Susanne
 APPLICANT: Polanyi-Bald, Laura
 APPLICANT: Prueser, Inken
 APPLICANT: Stehr, Regina
 APPLICANT: Maurer, Karl-Heinz
 TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PEROXYDOLASE ACTIVITY
 FILE REFERENCE: HENK-0134 / H5698
 CURRENT APPLICATION NUMBER: US/11/156,062
 CURRENT FILING DATE: 2005-06-17
 PRIOR APPLICATION NUMBER: PCT/EP2003/014127
 PRIOR FILING DATE: 2003-12-20
 PRIOR APPLICATION NUMBER: DE 102 60 903.9
 PRIOR FILING DATE: 2002-12-20
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.3
 SEQ ID NO 4
 LENGTH: 274
 TYPE: PRT
 ORGANISM: Artificial
 FEATURE:
 OTHER INFORMATION: Synthetic Construct
 US-11-156-062-4

Query Match 7.3%; Score 222.5; DB 7; Length 274;
 Best Local Similarity 29.9%; Pred. No. 3.5e-10;
 Matches 84; Conservative 35; Mismatches 109; Indels 53; Gaps 11;

QY 213 IVKADVAOSSGYLGQGIYAVADTGLDTRGSDSMHEAFRGKITALYALGRTNNANDTN 272
 Db 10 LIRKADKVOAQ-GFGKANVKAVALDTGLOASHPDLNVGG-----ASFVAGEAYN-TDGN 61
 QY 273 GHGTHVAGSV--LNGXTNKGMAPQANLVFQSIIMDSXGGLGSLPSNLQTLFSSQXSGAR 330
 Db 62 GHGTHVAGTVAALNTTGTGVLGAPSVSLYAVKVNSSG--SGSYSGIVSGIEMATTNGMD 119

```
QY 331 IHTNSGAVNGAYTTDSRNVDDYRKNDMTILFAAGNEXPNCG--GTISAPGTAKNATV 388
DB 120 VINMSLGGA---SGSTAKQAVDNAYARGVVVAAAGSSGSGNTNTTIGYPAKYDSVLAV 176
QY 389 GATEMLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPGTXILSARSSLAPDSFMA 448
DB 177 GA-----VDSNSNRASFSSVG-----ALELWMAFGAGVYSTYPT----- 210
QY 449 NHDSKYAYMGTSMATPIYAGNVA-----QLREHFVKNR 482
DB 211 ---NTYATWDTGTSMAFPHVAGAAALILSKHPILSASQVKNR 248

RESULT 14
US-11-156-062-10
; Sequence 10, Application US/11156062
; Publication No. US2005028173A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HEK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; PRIOR FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-10

Query Match 7.2%; Score 221.5; DB 7; Length 274;
Best Local Similarity 29.9%; Pred. No. 4.3e-10;
Matches 84; Conservative 35; Mismatches 109; Indels 53; Gaps 11;

QY 213 IYKADVAOSYGLYQCGQIVAVADTGLDTRGNDSSMHEAFRGKITALYALGRTNNANDTN 272
DB 10 LIKADKVAQ-GFKANVAVAVLDGTIGQASHDLDLVVGG-----ASFVAGSAYNA-DGN 61
QY 273 GHGTHVAGSV--LNGKXTNKGMAPOANLVFOSIMDSXGGLGGLPSNLQTLFSGQAXSAGAR 330
DB 62 GHGTHVAGTVALDHTTGVLAGVAPSVSSYAVKVLNMSG--SGSYSGIVSGIEMATTNGMD 119
QY 331 IHTNSGAAVNGAYTTDSRNVDDYRKNDMTILFAAGNEXPNCG--GTISAPGTAKNATV 388
DB 120 VINMSLGGA---SGSTAKQAVDNAYARGVVVAAAGSSGSGNTNTTIGYPAKYDSVLAV 176
QY 389 GATEMLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPGTXILSARSSLAPDSFMA 448
DB 177 GA-----VDSNSNRASFSSVG-----ALELWMAFGAGVYSTYPT----- 210
QY 449 NHDSKYAYMGTSMATPIYAGNVA-----QLREHFVKNR 482
DB 211 ---NTYATWDTGTSMAFPHVAGAAALILSKHPILSASQVKNR 248

RESULT 15
US-11-156-062-6
; Sequence 6, Application US/11156062
; Publication No. US2005028173A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne
```

```
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HEK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; PRIOR FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-6

Query Match 7.2%; Score 220.5; DB 7; Length 274;
Best Local Similarity 29.9%; Pred. No. 5.1e-10;
Matches 84; Conservative 35; Mismatches 109; Indels 53; Gaps 11;

QY 213 IYKADVAOSYGLYQCGQIVAVADTGLDTRGNDSSMHEAFRGKITALYALGRTNNANDTN 272
DB 10 LIKADKVAQ-GFKANVAVAVLDGTIGQASHDLDLVVGG-----ASFVAGSAYNA-DGN 61
QY 273 GHGTHVAGSV--LNGKXTNKGMAPOANLVFOSIMDSXGGLGGLPSNLQTLFSGQAXSAGAR 330
DB 62 GHGTHVAGTVALDHTTGVLAGVAPSVSSYAVKVLNMSG--SGSYSGIVSGIEMATTNGMD 119
QY 331 IHTNSGAAVNGAYTTDSRNVDDYRKNDMTILFAAGNEXPNCG--GTISAPGTAKNATV 388
DB 120 VINMSLGGA---SGSTAKQAVDNAYARGVVVAAAGSSGSGNTNTTIGYPAKYDSVLAV 176
QY 389 GATEMLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPGTXILSARSSLAPDSFMA 448
DB 177 GA-----VDSNSNRASFSSVG-----ALELWMAFGAGVYSTYPT----- 210
QY 449 NHDSKYAYMGTSMATPIYAGNVA-----QLREHFVKNR 482
DB 211 ---NTYATWDTGTSMAFPHVAGAAALILSKHPILSASQVKNR 248

Search completed: April 7, 2006, 18:04:21
Job time : 18.5129 secs
```


This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 00:56:02 ; Search time 9406.15 Seconds
(without alignment) 11602.984 Million cell updates/sec

Title: US-10-784-870-3
Perfect score: 1920
Sequence: 1 atgagagaagaagaaggtgtt.....cgttgcaattgtgaactaa 1920

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pal:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_scs:*
11: gb_ey:*
12: gb_un:*
13: gb_vl:*
14: gb_hlg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1920	100.0	1920	1	AB046403 Bacillus
2	1920	100.0	1920	6	AR368116 Sequence
3	1920	100.0	1920	6	AR562464 Sequence
4	1478.8	77.0	1923	1	AB084155 Bacillus
5	1477.2	76.9	1923	6	AB051423 Bacillus
6	1477.2	76.9	1923	6	CO888285 Sequence
7	1477.2	76.9	1923	6	CO889309 Sequence
8	1477.2	76.9	1923	6	AR368117 Sequence
9	1477.2	76.9	1923	6	AR368118 Sequence
10	1477.2	76.9	1923	6	AR562465 Sequence
11	1477.2	76.9	1923	6	AR562466 Sequence
12	1136.8	59.2	3003	6	BD062155 Nucleic a
13	1136.8	59.2	3003	6	BD062155 Nucleic a
14	1002.6	52.2	1305	6	AB039476 Sequence
15	930	48.4	2218	6	E03808 DNA encodin
16	823.6	38.9	1302	1	AB046406 Bacillus
17	747.6	38.9	1299	1	AB046404 Bacillus
18	738	38.4	1299	1	AB046405 Bacillus

19	734.8	38.3	1299	1	AB046402 Bacillus
20	89	4.6	346148	1	BX842552 Bellowib
21	58.2	3.0	7218	6	I6494 Sequence 14
22	47.4	2.5	2000	6	AX655393 Sequence
23	47	2.4	2000	6	AX655393 Sequence
24	46	2.4	5772	2	AF466309 Dicyoste
25	45.2	2.4	110000	1	BA000004_20
26	44.6	2.3	250541	1	AE017004 Bacillus
27	44.2	2.3	125020	8	AF429315 Homo sapi
28	44	2.3	250029	2	AE014820 Plasmodiu
29	43.6	2.3	6115	2	DDU20432 Dicyoste
30	43.6	2.3	110000	1	AE017225_18
31	43.6	2.3	110000	1	AE017334_18
32	43.6	2.3	110000	1	AE017355_19
33	43.6	2.3	110000	1	CP000001_19
34	43.6	2.3	291490	1	AE017030 Bacillus
35	43.4	2.3	110000	1	AE017180_35
36	43.4	2.3	154712	14	AC139734 Sus scro
37	43.4	2.3	211431	14	AC111594 Ratrus no
38	43.4	2.3	246852	14	AC096194 Ratrus no
39	43.4	2.3	255459	14	AC127743 Ratrus no
40	42	2.2	110000	1	TANNA_03 Continuation (4 of
41	41.2	2.1	110000	1	BA000028_24 Continuation (25 o
42	41	2.1	1141	6	AR579680 Sequence
43	41	2.1	1141	6	AX083744 Sequence
44	41	2.1	4839	1	VIBPSA19 D85144 Vibrio para
45	41	2.1	168438	14	AC151436 Carolina

ALIGNMENTS

RESULT 1	AB046403	1920 bp	DNA	linear	BCT 10-MAY-2002
LOCUS	AB046403	Bacillus sp. 9860	PROA gene for protease, complete cds.		
DEFINITION	AB046403				
ACCESSION	AB046403.2	GI:20521152			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Bacillus sp. 9860				
REFERENCE	Bacillus sp. 9860				
AUTHORS	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
TITLE	Saeki, K., Okuda, M., Hatada, Y., Kobayashi, T., Ito, S., Takami, H. and				
JOURNAL	Horikoshi, K.				
PUBMED	Novel oxidatively stable subtilisin-like serine proteases from				
REFERENCES	alkaliphilic Bacillus spp.: enzymatic properties, sequences, and				
TITLE	evolutionary relationships				
JOURNAL	Biochem. Biophys. Res. Commun. 279 (2), 313-319 (2000)				
COMMENT	Submitted (20-JUN-2000) Katsuhisa Saeki, Kao corporation,				
FEATURES	Biological Science Laboratory, Ichikaimachi Akabane 2606, Hega,				
SOURCE	Tohichi 321-3497, Japan (E-mail:387185@katsanet.kao.co.jp,				
gene	Tel:81-285-68-7400, Fax:81-285-68-7403)				
CDS	On May 9, 2002 this sequence version replaced gi:12381938.				
	Location/Qualifiers				
	1..1920				
	/organism="Bacillus sp. 9860"				
	/mol_type="genomic DNA"				
	/strain="9860"				
	/db_xref="taxon:133778"				
	1..1920				
	/gene="PROA"				
	/gene="PROA"				
	/codon_start=1				
	/transl_table=11				
	/product="protease"				
	/protein_id="BAB21266.2"				
	/db_xref="GI:20521155"				

/translation="MRKKVPLSVLSAAIIITVLANNSAGADATPDLDKIGIOTT
DVSGRSKOROTGAAAFVYESENVKILKELKLETPVANKLHIYQNPPIILBEKOK
LETTAKLIDYIPEVAYIIVEYEGDVQSVRSIEHVESYEDLPKIDPQJTGASST
LVKALADPTKONKNEVORLGEIEIAQVYASNDVHIITKPEKWNDAKRIVADVA
OSSVGLYGOQOI VAVADTGLDTRNDS SNEHAFRGI TALALGSTNNANDNGHT
VAGSVLGAITNKGA POANTVFOI IMOSGGLGPNLQTLFEOASAPARHITNS
MGAIVNGAYTTDSNRVDYVRKNDMTILFAAGNERPNNGTIL SAPETNAITLYGTEN
LRPSGAYADININHVADPSRSRPTKDGKIKPDVNAPOGYIISASBSLA PDSPFANHD
SKIVMGSTMAATPIVAGNVAQLREHFVYKNGIIPKPSLKAALINAGADVGLPYNG
NOGMRVTLIDKSLINAVYNESALSTSKATYTFATAGKPLKISLVNSDAPASTYAS
VTLVNDLDELVIAPNGTRYVGNDF SAPEDNMDGNHNVENVFINSPOSTYTIIEVOAY
NPVQPNFSAIYV"

ORIGIN

v

Query Match 100.0%; Score 1920; DB 1; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAAAGAAAGAGTGTGTTTATCTGTTTATCAGCTGCAGCGATTCCTGCACTGT 60
DB 1 ATGAGAAAGAAAGAGTGTGTTTATCTGTTTATCAGCTGCAGCGATTCCTGCACTGT 60
QY 61 GCATTAAACATCCCTCGGCTGTGTATGCAAGCACTTTGATCTGGAATTTTAAAGAAAT 120
DB 61 GCATTAAACATCCCTCGGCTGTGTATGCAAGCACTTTGATCTGGAATTTTAAAGAAAT 120
QY 121 CAAACAAACCCGATGTCACTGTTTCTTCGAAACAGCGACAAACAGGTGCGGCTGCATT 180
DB 121 CAAACAAACCCGATGTCACTGTTTCTTCGAAACAGCGACAAACAGGTGCGGCTGCATT 180
QY 181 CTGGTGAAGTCTGAAATGTGAAACCTCTTAAAGATTCCTAAGAACTTGAAACAGTA 240
DB 181 CTGGTGAAGTCTGAAATGTGAAACCTCTTAAAGATTCCTAAGAACTTGAAACAGTA 240
QY 241 CCGGCAAAATATTAATCTCCATATGTGCCAATTCAATGCGCCCATTTTAAAGAAACAAA 300
DB 241 CCGGCAAAATATTAATCTCCATATGTGCCAATTCAATGCGCCCATTTTAAAGAAACAAA 300
QY 301 CAGAAAGCTAAGACAACTGAGCAAAAGATTCCTGACTACATCCCTGATTAATGCAATAT 360
DB 301 CAGAAAGCTAAGACAACTGAGCAAAAGATTCCTGACTACATCCCTGATTAATGCAATAT 360
QY 361 GTGAGTATGAGGGGGAATGTTCACTCAAAAGTCCGCTCCATTGAAACAGGTGAAATCAGTG 420
DB 361 GTGAGTATGAGGGGGAATGTTCACTCAAAAGTCCGCTCCATTGAAACAGGTGAAATCAGTG 420
QY 421 GAGCCATACCTGCGCAAAATACAAATAGATCCCAAGCTTTTCAAAAAGGCGCATCGACG 480
DB 421 GAGCCATACCTGCGCAAAATACAAATAGATCCCAAGCTTTTCAAAAAGGCGCATCGACG 480
QY 481 CTGGTGAAGCTTGGCCCTTGATACGAAGCAAAACAAATAAAGATGCAATTAAGAGGC 540
DB 481 CTGGTGAAGCTTGGCCCTTGATACGAAGCAAAACAAATAAAGATGCAATTAAGAGGC 540
QY 541 ATCGAGAAATGCTCAGTACGTAGCAAGCAATGACGTCAATTATTTACGCAAGCT 600
DB 541 ATCGAGAAATGCTCAGTACGTAGCAAGCAATGACGTCAATTATTTACGCAAGCT 600
QY 601 GAAATTAAGGTATGATGATGTGGCCAGAGGTAATGTCAAAAGCGAGATGTGCACAGAGC 660
DB 601 GAAATTAAGGTATGATGATGTGGCCAGAGGTAATGTCAAAAGCGAGATGTGCACAGAGC 660
QY 661 AGCTACGGTTGTATGAGCAAGGCCAGATTTGCGAGTTGCCGATACGGAATTTGATACA 720
DB 661 AGCTACGGTTGTATGAGCAAGGCCAGATTTGCGAGTTGCCGATACGGAATTTGATACA 720
QY 721 GGAAGAAACGACAGTTCAGTACGAGCCTTCGCGGTAAATTAACAGCACTATATGCA 780
DB 721 GGAAGAAACGACAGTTCAGTACGAGCCTTCGCGGTAAATTAACAGCACTATATGCA 780
QY 781 CTGGGTCCGACGAATATGCGAATGACGAACGCTCATGTGATCCCATGTGCAAGTTCCG 840
DB 781 CTGGGTCCGACGAATATGCGAATGACGAACGCTCATGTGATCCCATGTGCAAGTTCCG 840

QY 841 GTATTAGAAATGCGCGCAACGAATTAAGGAATGGCACTCAAGCGAATCTGGTTTTTCAA 900
DB 841 GTATTAGAAATGCGCGCAACGAATTAAGGAATGGCACTCAAGCGAATCTGGTTTTTCAA 900
QY 901 TCCATCATGATATGACAGTGTGAGGCTTGAGAGCTTGCCTTCCAACTTGCAACCTTATTC 960
DB 901 TCCATCATGATATGACAGTGTGAGGCTTGAGAGCTTGCCTTCCAACTTGCAACCTTATTC 960
QY 961 AGCCAGCATTCAGTGCAGGTGCCAGAAATTCATACAACTCTGAGGGGCGAGCGGTGAAT 1020
DB 961 AGCCAGCATTCAGTGCAGGTGCCAGAAATTCATACAACTCTGAGGGGCGAGCGGTGAAT 1020
QY 1021 GGGGCTTACAGACAGATTTCCAGAAATGTGATGACTATGPAAGAAATGATATGACG 1080
DB 1021 GGGGCTTACAGACAGATTTCCAGAAATGTGATGACTATGPAAGAAATGATATGACG 1080
QY 1081 ATTCCTTTCGCGGCTGGAATGAAGGCCGAAACCGGCTACCATCATGTCACCTGTACG 1140
DB 1081 ATTCCTTTCGCGGCTGGAATGAAGGCCGAAACCGGCTACCATCATGTCACCTGTACG 1140
QY 1141 GCTAATAACGCCATTAACAGTGGCGCAACGAAACCTGCTCCAGCTTGGTTCCTAT 1200
DB 1141 GCTAATAACGCCATTAACAGTGGCGCAACGAAACCTGCTCCAGCTTGGTTCCTAT 1200
QY 1201 GCAGTAAATTAACAGAGTTCGACAGTTCCTCCGCGGCGCAAAAGATGGGCGA 1260
DB 1201 GCAGTAAATTAACAGAGTTCGACAGTTCCTCCGCGGCGCAAAAGATGGGCGA 1260
QY 1261 ATCAAGCTGATGTATGCGCCAGGACATACATTTTATGAGCAAGATCTTCTTGCA 1320
DB 1261 ATCAAGCTGATGTATGCGCCAGGACATACATTTTATGAGCAAGATCTTCTTGCA 1320
QY 1321 CCGGATTCCTCTTCTGCGGCGAATCATGACGCAATATGCTATATGGGTGGAACCTTC 1380
DB 1321 CCGGATTCCTCTTCTGCGGCGAATCATGACGCAATATGCTATATGGGTGGAACCTTC 1380
QY 1381 ATGGCAACCCGATTTGTTGCGGGGAATGTTGACAGCTCCGTGAGCATTTTGTGAAAT 1440
DB 1381 ATGGCAACCCGATTTGTTGCGGGGAATGTTGACAGCTCCGTGAGCATTTTGTGAAAT 1440
QY 1441 AGAGGAATCACTCTTAAGCCTTCCCTATTTGAAGACGTTTGAATTCAGGTGCTGAT 1500
DB 1441 AGAGGAATCACTCTTAAGCCTTCCCTATTTGAAGACGTTTGAATTCAGGTGCTGAT 1500
QY 1501 GTTGATTTGGGTATCCGAACGGAACCAAGATGGGCGAGTGAACCTTGATTAATCG 1560
DB 1501 GTTGATTTGGGTATCCGAACGGAACCAAGATGGGCGAGTGAACCTTGATTAATCG 1560
QY 1561 TTGAACGTTGCTATGTGAACGAATCCAGTCCCTATCACTAGCCAAAGACGACATAT 1620
DB 1561 TTGAACGTTGCTATGTGAACGAATCCAGTCCCTATCACTAGCCAAAGACGACATAT 1620
QY 1621 ACCTTACTGACAGCGCGGCAAGCATTTGAAATCTCCTGTATGTGTCGATGCCCT 1680
DB 1621 ACCTTACTGACAGCGCGGCAAGCATTTGAAATCTCCTGTATGTGTCGATGCCCT 1680
QY 1681 GCAAGCACTAATGCTTCTGTAACTCCTGTGCAATGATTTGGATTTGTCATTAACGACCA 1740
DB 1681 GCAAGCACTAATGCTTCTGTAACTCCTGTGCAATGATTTGGATTTGTCATTAACGACCA 1740
QY 1741 AACGGAACAATATGTGGGAATGACTTCTCAGACCACTTGAACAATACGTGATGGC 1800
DB 1741 AACGGAACAATATGTGGGAATGACTTCTCAGACCACTTGAACAATACGTGATGGC 1800
QY 1801 CGCAATTAACGTAAGAAATGATTTATTAATTCGCCCAAGTGAACATATACATTAG 1860
DB 1801 CGCAATTAACGTAAGAAATGATTTATTAATTCGCCCAAGTGAACATATACATTAG 1860
QY 1861 GTGCAAGATTAATATGTCGGTGGACCAAAAATTTCTGTTGGCAATTTGTAACATA 1920
DB 1861 GTGCAAGATTAATATGTCGGTGGACCAAAAATTTCTGTTGGCAATTTGTAACATA 1920

RESULT 2
AR368116
LOCUS AR368116 1920 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 3 from patent US 6376227.
ACCESSION AR368116
VERSION AR368116.1 GI:34601777
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
1 (bases 1 to 1920)
AUTHORS Takaiwa, M., Okuda, M., Saeki, K., Kubota, H., Hitomi, J., Kageyama, Y.,
Shikata, S. and Nonura, M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 3 23-APR-2002;
Kao Corporation; Tokyo;
JPX;
FEATURES
source 1..1920
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1920; DB 6; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAAAGAGAGAGTGTGTTTATCTGTTTATCAGCTGCAAGCAATCTGTCAGTGT 60
DB 1 ATGAGAAAGAGAGAGTGTGTTTATCTGTTTATCAGCTGCAAGCAATCTGTCAGTGT 60
QY 61 GATTTAAACAATCCCTCGGCTGTGTATGCAAGGACTTTGATCTGATTTTAAAGAAAT 120
DB 61 GATTTAAACAATCCCTCGGCTGTGTATGCAAGGACTTTGATCTGATTTTAAAGAAAT 120
QY 61 GATTAAACAATCCCTCGGCTGTGTATGCAAGGACTTTGATCTGATTTTAAAGAAAT 120
DB 61 GATTAAACAATCCCTCGGCTGTGTATGCAAGGACTTTGATCTGATTTTAAAGAAAT 120
QY 121 CAACAACAACGATGTCAAGTGTCTCCAAACAGCGACAACAGGTCGGCTGCATT 180
DB 121 CAACAACAACGATGTCAAGTGTCTCCAAACAGCGACAACAGGTCGGCTGCATT 180
QY 121 CAAACAACAACGATGTCAAGTGTCTCCAAACAGCGACAACAGGTCGGCTGCATT 180
DB 121 CAAACAACAACGATGTCAAGTGTCTCCAAACAGCGACAACAGGTCGGCTGCATT 180
QY 181 CTGTGTAGAGTCTGAAAATGTGAAACTTCTTAAAGATTCCTAAGAACTTGAACAAGTA 240
DB 181 CTGTGTAGAGTCTGAAAATGTGAAACTTCTTAAAGATTCCTAAGAACTTGAACAAGTA 240
QY 181 CTGTGTAGAGTCTGAAAATGTGAAACTTCTTAAAGATTCCTAAGAACTTGAACAAGTA 240
DB 181 CTGTGTAGAGTCTGAAAATGTGAAACTTCTTAAAGATTCCTAAGAACTTGAACAAGTA 240
QY 241 CCGGCAAAATTAACCTCCATATTTGTCATTCATTCATGCCCCCATTTTAAAGAAACAAA 300
DB 241 CCGGCAAAATTAACCTCCATATTTGTCATTCATTCATGCCCCCATTTTAAAGAAACAAA 300
QY 301 CAGAAGCTAGAGCAACTGTGAGCAAAAGATTCTGACTATCCCTGATTATGATATATT 360
DB 301 CAGAAGCTAGAGCAACTGTGAGCAAAAGATTCTGACTATCCCTGATTATGATATATT 360
QY 361 GTGAGTATGAGGGGGATGTTGATGTCAAAGTCCGCTCCATTTGAACAAGTGAATCAGTG 420
DB 361 GTGAGTATGAGGGGGATGTTGATGTCAAAGTCCGCTCCATTTGAACAAGTGAATCAGTG 420
QY 421 GAGCCATACCTGCGCAAAATCAAAATAGATCCCAAGCTTTTCAAAAAGGCCATGACG 480
DB 421 GAGCCATACCTGCGCAAAATCAAAATAGATCCCAAGCTTTTCAAAAAGGCCATGACG 480
QY 481 CTGTGAAAGCGTTGCGCTGATAGCAAGGCAAGCAATAAAGATGTGAATTAAGAGGC 540
DB 481 CTGTGAAAGCGTTGCGCTGATAGCAAGGCAAGCAATAAAGATGTGAATTAAGAGGC 540
QY 541 ATCGAGAAATCGCTCAGTACGTAGCAAGCAATGACGTTCATTATTTACGCAAAAGCT 600
DB 541 ATCGAGAAATCGCTCAGTACGTAGCAAGCAATGACGTTCATTATTTACGCAAAAGCT 600
QY 601 GAATTTAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 GAATTTAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 AGCTACGGTTTGTATGACAAAGCCAGATTTGTGCGAGTTGCCGATACCTGATTTGATACA 720

DB 661 AGCTACGGTTTGTATGACAAAGCCAGATTTGTGCGAGTTGCCGATACCTGATTTGATACA 720
QY 721 GGAAGAAACGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 721 GGAAGAAACGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 781 CTGGGTCCGACGAATTAATGCGAATGATGATGATGATGATGATGATGATGATGATGATGATG 840
DB 781 CTGGGTCCGACGAATTAATGCGAATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 841 GTATTAGAAATGCGCAACGAATTAAGAAATGAGCACTCAAGCGAATCTGTTTTCAA 900
DB 841 GTATTAGAAATGCGCAACGAATTAAGAAATGAGCACTCAAGCGAATCTGTTTTCAA 900
QY 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
DB 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 961 ACCCAAGCAATTCAGATGCAAGTGTCCAGAAATTCATCAAACTCTGGGGGGGCAAGCGTGAT 1020
DB 961 ACCCAAGCAATTCAGATGCAAGTGTCCAGAAATTCATCAAACTCTGGGGGGGCAAGCGTGAT 1020
QY 1021 GGGGCTTACACGACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB 1021 GGGGCTTACACGACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 ATTCTTTTCCGCGCTGGGAGATGAAAGCGCGAAGCGGATACATGATGATGATGATGATGATG 1140
DB 1081 ATTCTTTTCCGCGCTGGGAGATGAAAGCGCGAAGCGGATACATGATGATGATGATGATGATG 1140
QY 1141 GCTAATAAACGCAATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
DB 1141 GCTAATAAACGCAATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1201 GCAAGTAATATTAACCAAGTGTGCAAGATTCCTGTCGGGGGGGCAAAAGATGAGGCGCA 1260
DB 1201 GCAAGTAATATTAACCAAGTGTGCAAGATTCCTGTCGGGGGGGCAAAAGATGAGGCGCA 1260
QY 1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
DB 1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
QY 1321 CCGGATTCCTCTTCTGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1321 CCGGATTCCTCTTCTGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
QY 1381 ATGGCAACACGGAATGTTGCGGGGGAATGTTGCAAGCTCCGAGACATTTGTGAATAAT 1440
DB 1381 ATGGCAACACGGAATGTTGCGGGGGAATGTTGCAAGCTCCGAGACATTTGTGAATAAT 1440
QY 1441 AGAGGAATCACTCTTAAGCTTCCCTATTTGAAGAAGCTTTGATGATGATGATGATGATGATG 1500
DB 1441 AGAGGAATCACTCTTAAGCTTCCCTATTTGAAGAAGCTTTGATGATGATGATGATGATGATG 1500
QY 1501 GTTGAATGAGGTTATCCGAACGGAACCAAGGATGAGGCGGAGTCACTCGATTAATATG 1560
DB 1501 GTTGAATGAGGTTATCCGAACGGAACCAAGGATGAGGCGGAGTCACTCGATTAATATG 1560
QY 1561 TTGAAGTGTGCTATGTTGAAGGAATCCAGTGCCTTATCAACATGCAAAAGCGACATAT 1620
DB 1561 TTGAAGTGTGCTATGTTGAAGGAATCCAGTGCCTTATCAACATGCAAAAGCGACATAT 1620
QY 1621 ACCTTTACTGCAACGGCGGGCAAGCAATTTGAATAATCTCTGATGATGATGATGATGATGATG 1680
DB 1621 ACCTTTACTGCAACGGCGGGCAAGCAATTTGAATAATCTCTGATGATGATGATGATGATGATG 1680
QY 1681 GCAAGCACTACGCTTCTGTTAATCCTGATCAATGATTTGATTTGATTTGATTTGATTTGATTTG 1740
DB 1681 GCAAGCACTACGCTTCTGTTAATCCTGATCAATGATTTGATTTGATTTGATTTGATTTGATTTG 1740
QY 1741 AACGGAACAAGTATGTTGGGGAATGATTTCCAGACCAATTTGAATTAACCTGAGATGAGC 1800

Db 1741 AACGGAACAAGATATGTCCGGAAAGACTTCTCAGACCACTTTGACATTAAGTGGATGAC 1800
Qy 1801 CGCAATTAACGTAAAGATATTTTATTTATTTGCCCCCAAGTGAACATATACCTTGG 1860
Db 1801 CGCAATTAACGTAAAGATATTTTATTTGCCCCCAAGTGAACATATACCTTGG 1860
Qy 1861 GTGCAAGCATTAATGTGCTCCGGTGGACCAAAACTTCTCGTTGGCAATTTGAACATTA 1920
Db 1861 GTGCAAGCATTAATGTGCTCCGGTGGACCAAAACTTCTCGTTGGCAATTTGAACATTA 1920

RESULT 3
AR562464
LOCUS AR562464 1920 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 3 from patent US 6759228.
ACCESSION AR562464
VERSION AR562464.1 GI:53976505
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1920)
AUTHORS Takeiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kagayama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6759228-A 3 06-JUL-2004;
Kao Corporation; Tokyo;
JPK;

FEATURES
source 1..1920
location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 1920; DB 6; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 541 ATGAGGAATATGCTCAGTACGTAGCAAGCAATGACGATTAATTAAGGCAAAAGCCT 600
Db 541 ATGAGGAATATGCTCAGTACGTAGCAAGCAATGACGATTAATTAAGGCAAAAGCCT 600
Qy 601 GAATATTAAGTATGATATGATGTGGCGAAGATTTGTCAAAAGCGATGTGGCAAGAGC 660
Db 601 GAATATTAAGTATGATATGATGTGGCGAAGATTTGTCAAAAGCGATGTGGCAAGAGC 660
Qy 661 AGCTACGTTTGTATGACAAAGCCAGATTGTCCGAGTTGCCGATACCTGATTTGATACA 720
Db 661 AGCTACGTTTGTATGACAAAGCCAGATTGTCCGAGTTGCCGATACCTGATTTGATACA 720
Qy 721 GGAAGAAACGACAGTTGATGATGAAAGCCTTCGCGGTAAATTAACAGCACTAATATCA 780
Db 721 GGAAGAAACGACAGTTGATGATGAAAGCCTTCGCGGTAAATTAACAGCACTAATATCA 780
Qy 781 CTGGGTCCGACGAATTAATGCGAATGATACGAACGGTCAATGTAACCAATGTGGCAGTTTCG 840
Db 781 CTGGGTCCGACGAATTAATGCGAATGATACGAACGGTCAATGTAACCAATGTGGCAGTTTCG 840
Qy 841 GTATTTAGAAATGCGCGCAACGAATTAAGAAATGCGACCTCAAGCGAAATCTGTTTTCAA 900
Db 841 GTATTTAGAAATGCGCGCAACGAATTAAGAAATGCGACCTCAAGCGAAATCTGTTTTCAA 900
Qy 901 TCCATCATGATGATACAGTGTGGGCTTGAAGGCTTGCCTTCATTTGCAAACTTAATTC 960
Db 901 TCCATCATGATGATACAGTGTGGGCTTGAAGGCTTGCCTTCATTTGCAAACTTAATTC 960
Qy 961 AGCCAAAGCATTCAGTGCAGGTGCGCAATTCATCAAACTCTCGGGGGCGACGGTGAAT 1020
Db 961 AGCCAAAGCATTCAGTGCAGGTGCGCAATTCATCAAACTCTCGGGGGCGACGGTGAAT 1020
Qy 1021 GGGGCTTACACGACAGATTCGAAATGTGATGATATGTAAGAAATGATATGACG 1080
Db 1021 GGGGCTTACACGACAGATTCGAAATGTGATGATATGTAAGAAATGATATGACG 1080
Qy 1081 ATTCTTTTCCGGGCGTGGAAATGAAGGCCGAACGGGCTTGCATATGCACTGCACTGTACG 1140
Db 1081 ATTCTTTTCCGGGCGTGGAAATGAAGGCCGAACGGGCTTGCATATGCACTGTACG 1140
Qy 1141 GCTAAATAACGCGCATACAGTGGCGCAACCGAAACCTGCTCCAGACTTGGTTCAT 1200
Db 1141 GCTAAATAACGCGCATACAGTGGCGCAACCGAAACCTGCTCCAGACTTGGTTCAT 1200
Qy 1201 GCAGATTAATTAACAAGTTGCAAGTTCCTTCCGTTGCGCGCAAAAGATGGGCGA 1260
Db 1201 GCAGATTAATTAACAAGTTGCAAGTTCCTTCCGTTGCGCGCAAAAGATGGGCGA 1260
Qy 1261 ATCAAGGCTGATGTCATGAGGCGCGGACATACATTTTATCAGAAATCTTCTTGA 1320
Db 1261 ATCAAGGCTGATGTCATGAGGCGCGGACATACATTTTATCAGAAATCTTCTTGA 1320
Qy 1321 CCCGATTCCTCTCTGCGGCGAATCATGACAGCAAAATATGCTTATATGGGTGAACGTC 1380
Db 1321 CCCGATTCCTCTCTGCGGCGAATCATGACAGCAAAATATGCTTATATGGGTGAACGTC 1380
Qy 1381 ATGSCAACACCGAATTTGTTGGGGGAATGTTGCAAGCTCCGTTGCAATTTTGTGA 1440
Db 1381 ATGSCAACACCGAATTTGTTGGGGGAATGTTGCAAGCTCCGTTGCAATTTTGTGA 1440
Qy 1441 AGAGGAATCACTCCTAAGCCTTCCCTATTTGAAGCAGCTTTGATGAGAGTGTGCTGAT 1500
Db 1441 AGAGGAATCACTCCTAAGCCTTCCCTATTTGAAGCAGCTTTGATGAGAGTGTGCTGAT 1500
Qy 1501 GTTGATTTGGGTATATCCGAACGAAACCAAGATGGGGCGAGTGAACCTGTGATAAATCG 1560
Db 1501 GTTGATTTGGGTATATCCGAACGAAACCAAGATGGGGCGAGTGAACCTGTGATAAATCG 1560
Qy 1561 TTGAACGTTGCTATATGTAAGCAATTCAGTGCCTTATCAACTAGCCAAAAGCGACATAT 1620
Db 1561 TTGAACGTTGCTATATGTAAGCAATTCAGTGCCTTATCAACTAGCCAAAAGCGACATAT 1620

QY 1621 ACCTTACTGCAACGGCGGAGCAATGAAATCTCCGTGTATGTGCGATGCGCCT 1680
DB 1621 ACCTTACTGCAACGGCGGAGCAATGAAATCTCCGTGTATGTGCGATGCGCCT 1680
QY 1681 GCAAGCACTACTGCTTCTGTAAACCTGTGTAAATGATTTGGTCAATTCAGACCA 1740
DB 1681 GCAAGCACTACTGCTTCTGTAAACCTGTGTAAATGATTTGGTCAATTCAGACCA 1740
QY 1741 AACGGAACAAGATGTGCGGAAATGACTTCTGACACCAATTTGACATTAATCGGATGCG 1800
DB 1741 AACGGAACAAGATGTGCGGAAATGACTTCTGACACCAATTTGACATTAATCGGATGCG 1800
QY 1801 GCGAATTAAGTAAATGATTTATTTATTTATGCCCCAAGTGAACATTAATCGATGAG 1860
DB 1801 GCGAATTAAGTAAATGATTTATTTATTTATTTGCCCCAAGTGAACATTAATCGATGAG 1860
QY 1861 GTGCAAGCATATATATGTGCGGTGAGCAACAAATCTTCTGTGCAATTTGAACTAA 1920
DB 1861 GTGCAAGCATATATATGTGCGGTGAGCAACAAATCTTCTGTGCAATTTGAACTAA 1920

RESULT 4
AB084155 1923 bp DNA linear BCT 02-SRP-2003
LOCUS AB084155
DEFINITION Bacillus sp. KSM-9865 gene for protease, complete cds.
ACCESSION AB084155
VERSION AB084155.1 GI:34392386
KEYWORDS
SOURCE
ORGANISM
Bacillus sp. KSM-9865
Bacillus sp. KSM-9865
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

REFERENCE
AUTHORS 1 Okuda, M., Saeki, K. and Kobayashi, T.
TITLE Bacillus sp. KSM-9865 protease gene
JOURNAL Published Only in Database (2003)
REFERENCE 2 (bases 1 to 1923)
AUTHORS Okuda, M., Saeki, K. and Kobayashi, T.
TITLE Direct Submission
JOURNAL Submitted (18-Apr-2002) Mitsuyoshi Okuda, Kao Corporation,
Biological Science Laboratory; Ichikaimachi Akabane 2606, Haga,
Tochigi 321-3497, Japan (E-mail:okuda.mitsuyoshi@kao.co.jp,
Tel:81-285-68-7543, Fax:81-285-68-7547)

FEATURES
Source
Location/Qualifiers
1..1923
/organism="Bacillus sp. KSM-9865"
/mol_type="genomic DNA"
/strain="KSM-9865"
/db_xref="taxon:192495"
1..1923
/codon_start=1
/transl_table=1
/product="protease"
/protein_id="BAC82522.1"
/db_xref="GI:34392387"
/translation="MRKKRYPLSTLSAAALISTVALSNPSAGARNFDLPKGIQTITDAGPSKQGTGAAATLVESNVLPGKLTQKLETVANNKLAHVQNGPLBETKQLEKGAFLIPYATVIVIEGVEKSAITIEBESVBPPIYRIDPOLFTFGASBLVAVADLTQNKKEVOLRGILQAGATSNVDVITAKIEGNNVANGIVADVAOSSYGLGGQIIVADTGLDTGRNDSMRAFRGKTALVYALGRTNANDNCHGTHTAGSVLGNGSTNGMAPOANLVPSINDSGGLGLSNQTLFSPQVSKARHTNHWGAIVKATITDSRNDVDYRKNDMTLLFPAENGEPNGGTSIATGKMAITVGATNLRPSGVAONINHVAPQSRGPTKDRIPEDVNAPOTFILASNSLAPSSVANHDSKAVMGSTMAATPIVAGNVAAQLRHFVKKRGITPKPSLLKALIAAADIIGLQYPNQNGWGRVTLDSKLVAAYVNSSLSLSTQKATVSTATAGKPLKISLVMSDAPASTTASVTLVNDLDLVITAPNGTQYGVNDPTSPPYNNMORNVENVFIAPSGSYTTEVQAYNPVGPOTPSLAIVN"

ORIGIN
Query Match 77.0%; Score 1478.8; DB 1; Length 1923;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 1642; Conservative 0; Mismatches 272; Indels 0; Gaps 0;

7 AAGAAAGAGGTGTTTATCTGTTTATCACTGACGAGGATTCGTGCACTGTGCAATTA 66

DB 10 AAGAAAGAGGTGTTTATCTGTTTATCACTGACGAGGATTCGTGCACTGTGCAATTA 69
QY 67 AACAAATCCCTCGGCTGTGATGCAAGGACTTTTGAATCTGGAATTTTAAAGAAATTCAAACA 126
DB 70 AGTAATTCATCTGCAAGGTGTGCAAGGAATTTTGAATCTGGAATTTTAAAGAAATTCAGACA 129
QY 127 ACAACGATGTGATGTTTCTTCAACAGCGACAAACAGGTGCGCTGCAATTTCTGTG 186
DB 130 ATATACGTATGCTAAAGTTTCTTCAAGCGGCGAGATGTGTCTGCTGCTCATTTCTGTG 189
QY 187 GAGTCGAAAAATGTAATCTTTTAAAGATTTGCTAAAGAACTTGAACATTAACCGGCA 246
DB 190 GATTCGAAAAATGTAATCTTTTAAAGATTTGCTAAAGAACTTGAACATTAACCGGCA 249
QY 247 AATAATTAATCTCATATTTGCTCAATTCATATGAGCCCATTTTGAAGAAACAAACAGAG 306
DB 250 AATAATTAATCTCATATTTGCTCAATTCATATGAGCCCATTTTGAAGAAACAAACAGAG 309
QY 307 CTAGAGACAACTGAGCAAGATTTCTGACATACATCCCTGATTAATGATATATGTCGAG 366
DB 310 CTAGAGAAAAACAGGGGCAAGATTTCTGACATACATCCGATTAATGATTAATGTCGAG 369
QY 367 TATGAGGGGGATGTTCACTCAAAAGTCCGCTCCATTTGAACAGGTGAATCACTGAGACCA 426
DB 370 TATGAGGGGGATGTTTAAAGTCAAGCAACAGCACTTGAAGAGGTGAATCCGTAGACCT 429
QY 427 TACTTCCGAAATACAAATATGATCCCAAGCTTTTCACAAAGGCGCATGACGCTGTG 486
DB 430 TATTTGCCGATATACGAATATGATCCCAAGCTTTTCACAAAGGCGCATGACGCTGTG 489
QY 487 AAAGGTTGGCGCTTGATACGAAGCAAGCAATTAAGATGCAATTAAGGCACTGAG 546
DB 490 AAAGCAAGTGGCGCTTGATACGAATATGATCCCAAGCTTTTCACAAAGGCGCATGACG 549
QY 547 GAAATCGCTCACTAGTACGTAAGCAAGATACGCTCATTAATTAACGCAAGCTGATAT 606
DB 550 CAATATGCAATTCGCAACAGCAAGATATGATATTAATTAACGCAAGCTGATAT 609
QY 607 AAGGTGATGAATGATGTGGCGAGAGGTATTTGCAAGCGAGTGTGCAAGAGCACTAC 666
DB 610 AAGGTGATGAATGATGTGTGCAAGTGTGCAAGCGAGTGTGCAAGAGCACTAC 669
QY 667 GATTGTATGACAAAGCGCAATTTGTGCAAGTGTGCGATGCGATACGATTTGATACAGAA 726
DB 670 GGGTGTATGACAAAGCAAGCAAGTGTGACGCTGCGATACAGAGGCTGATACAGGTGCG 729
QY 727 AACGCACTTGTGATGCAAGAGCTTCCGCGGTAAATTAACAGCACTATATGCACTGGGT 786
DB 730 AATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 789
QY 787 CGAGCAATTAATGCAATGATACGAACGATACGATACGATACGATACGATACGATACGAT 846
DB 790 CGAGCAATTAATGCAATGATACGAATGATACGATACGATACGATACGATACGATACGAT 849
QY 847 GAAATGCGCAACGAATTAAGATGCACTCAAGCGAACTGTGTTTCAATTCATC 906
DB 850 GAAATGCGCTCACTCAATTAAGATGCACTCAAGCGAACTGTGTTTCAATTCATC 909
QY 907 ATGATACGATGCTGTGCGCTTGAAGGCTTGCCTTCAATTCGCAACCTTATTCAGCCAA 966
DB 910 ATGATACGATGCTGTGCGCTTGAAGGCTTGCCTTCAATTCGCAACCTTATTCAGCCAA 969
QY 967 GCAATTCAGTGAAGGCGCAAGATTCATCAAACTCTGAGGGGGCGAGGTGAATGGGGCC 1026
DB 970 GCAATTCAGTGTGCTGCTGCGCAAGATTCATCAAACTCTGAGGGGGCGAGGTGAATGGGGCC 1029
QY 1027 TACACGACAGATTCAGAAATGTGATGATCATATGTAAGAAATATGATATGCAATTCCT 1086
DB 1030 TACACGACAGATTCAGAAATGTGATGATCATATGTAAGAAATATGATATGCAATTCCT 1089
QY 1087 TTCGCGGCTGGGAATGAAGCGGCAAGCGGCTACATCATGTCACCTGTGATCGGCTAA 1146

Db 1090 TTCGCTGCGGGAATGAAGACCGGACCATCATGTCACGACAGCTAA 1149
 Qy 1147 AACCCATTAACAGTCGCGCAACCGAAACCTGCGTCAAGCTTGGTTCTATGACAT 1206
 Db 1150 AATGCAATTAACAGTCGCGCAACCGAAACCTGCGTCAAGCTTGGTTCTATGACAT 1209
 Qy 1207 AATTAATTAACAGTCGCGCAACCGAAACCTGCGTCAAGCTTGGTTCTATGACAT 1266
 Db 1210 AATTAATTAACAGTCGCGCAACCGAAACCTGCGTCAAGCTTGGTTCTATGACAT 1269
 Qy 1267 CCTGATGTCATGCGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1326
 Db 1270 CCGATGTCATGCGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1329
 Qy 1327 TCTCTCTTCTGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1386
 Db 1330 TCTCTCTTCTGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1389
 Qy 1387 ACACCGATTTGTCGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1446
 Db 1390 ACACCGATTTGTCGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1449
 Qy 1447 ATCACTCTTAAGCTTCTCTATTAAGACGCTTGAATGTCAGGTGCTGATGTTGGA 1506
 Db 1450 ATCACTCTTAAGCTTCTCTATTAAGACGCTTGAATGTCAGGTGCTGATGTTGGA 1509
 Qy 1507 TTGGGTTATCCGAAACGGAACCAAGATGCGGCGGATGACCTTGATTAATCGTTGAC 1566
 Db 1510 CTGGGTTATCCGAAACGGAACCAAGATGCGGCGGATGACCTTGATTAATCGTTGAC 1569
 Qy 1567 GTTGGTTATCCGAAACGGAACCAAGATGCGGCGGATGACCTTGATTAATCGTTGAC 1626
 Db 1570 GTTGGTTATCCGAAACGGAACCAAGATGCGGCGGATGACCTTGATTAATCGTTGAC 1629
 Qy 1627 ACTGCAACGCGGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1686
 Db 1630 ACTGCAACGCGGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1689
 Qy 1687 ACTACTGCTTCTGTAACCTTGTGTAATGATTTGATTTGATTAATGACCAACCAACGGA 1746
 Db 1690 ACTACTGCTTCTGTAACCTTGTGTAATGATTTGATTTGATTTGATTAATGACCAACGGA 1749
 Qy 1747 ACAAGATTTGTCGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1806
 Db 1750 ACAAGATTTGTCGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1809
 Qy 1807 AACGTAATTAATGATTTATTAATGTCGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1866
 Db 1810 AACGTAATTAATGATTTATTAATGTCGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1869
 Qy 1867 GCATTAATTAATGTCGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1920
 Db 1870 GCATTAATTAATGTCGCGGACCATCATGTTTATGACAAAGTCTTCTTTCGACCCGAT 1923
 RESULT 5
 AB051423 1923 bp DNA linear BCT 10-MAY-2002
 LOCUS Bacillus sp. KP43 PROF gene for protease, complete cds.
 DEFINITION AB051423
 ACCESSION AB051423
 VERSION AB051423.2 GI:20521154
 KEYWORDS
 SOURCE
 ORGANISM
 Bacillus sp. KSM-KP43
 Bacillus sp. KSM-KP43
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE
 AUTHORS Itoh, S. and Saeki, K.
 TITLE new protease
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1923)
 AUTHORS Saeki, K.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2000) Katsuhisa Saeki, KAO, CORPORATION,

2606,AKABANE,ICHIKAWACHI, HAGA, TOCHIGI 321-3486, Japan
 (E-mail:387185@katsanet.kao.co.jp, Tel:81285687471 (ex.7471),
 Fax:81285687403)
 On May 9, 2002 this sequence version replaced gi:14164344.
 COMMENT location/Qualifiers
 FEATURES
 SOURCE
 1..1923
 /organism="Bacillus sp. KSM-KP43"
 /mol_type="genomic DNA"
 /strain="KP43"
 /db_xref="taxon:109322"
 1..1923
 /gene="PROF"
 1..1923
 /gene="PROF"
 /codon_start=1
 /transl_table=1
 /product="protease"
 /protein_id="BAB5674.2"
 /db_xref="GI:20521155"
 /translation="MRKKKVFPLSVSAAAILSTVALSNPSGARRPDDPKIQT
 TDAKGFSGQQTGAALPLVSENVKLPKGLQKLETPANNKLIHIDPGLLETKQ
 QLEKTVGAKLIDYIPDYVIVEBDDVSAISTIEHVSVEPYLPIYIDQLFTKGS
 ELVAVALDTRKOKKEVOLRIGIQLAFLSNDVLYTAKPEYKMDVARIVKADV
 AQSRYGLGGQIVAVADTGLDGRNDSHBEAPRGIITLYALGRNNANDTGHGT
 HUGSYVGNSTKKGMAPOANIVPOSIMDGGGIGGIPSNLOTLPSQAVAGARLHIN
 SWGAIVGATTTDSRVNDYVRKNDITILPAAGREGNGCTISAPGTAKALITGATE
 NLRPSFQSVADNINHVAFSSKPTKGRIRKPDMAVGTILSKRSSLDPSFWAH
 DSKRVAVGGTSMAPPIVAGVADLRHFVRKRGITPPRSLKALIGADIGGYRN
 GNOGMRVLTDKSLNVAVNESLSISTQKATVSTFATACKPLKLSIWSAPASSTTA
 SVTLVNDLDELVTAPNGTQYVNDFTSPYNDMDGRNNEVFIINAPSGTYTIEVQA
 YNPVVGQTSIAIVN"
 ORIGIN
 Query Match 76.9%; Score 1477.2; DB 1; Length 1923;
 Best Local Similarity 85.7%; Pred. No. 0;
 Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
 Qy 7 AAGAAAGAGTGTGTTTATGCTGTTTATGAGTCGACGAGATTCGTGACGTGTCATTA 66
 Db 10 AAGAAAGAGTGTGTTTATGCTGTTTATGAGTCGACGAGATTCGTGACGTGTCATTA 69
 Qy 67 AACATCCCTGCGTGGTGAAGGACCTTTGATCGATTTGATTAAGAAATCAACA 126
 Db 70 AGTATCATTTGACAGGTGTCGACGAAATTTGATCGATTTGATTAAGAAATCAACA 129
 Qy 127 AACACGATGTCAGTGTGTTCTCCAAACAGGCAACAAAGTGGCGTGCATTTCTGAGT 186
 Db 130 AACATGATGTCAGTGTGTTCTCCAAACAGGCGACAGCTGGTGTGCTGTTTCTGAGT 189
 Qy 187 GAGTCTGAAATGCGAACTTCTTAAGGATTCGTAAGAACTTGAACAGTACCGGCA 246
 Db 190 GAACTGAAATGCGAACTTCTCCAAAGGTTTGAAGAAAGCTTGAACAGTACCGGCA 249
 Qy 247 AATTAATAACCTCATATTGTCATATGATGAGCCCATTTTGAAGAAACAAACAGAA 306
 Db 250 AATTAATAACCTCATATTGTCATATGATGAGCCCATTTTGAAGAAACAAACAGAA 309
 Qy 307 CTAGAGCAACTGAGCAAAAGATTCTGACTACATCCCTGATTAATGATATATTGTCAG 366
 Db 310 CTGAGAAAACAGGCGGCAAGATTCTGACTACATCTGATTAATGATATATTGTCAG 369
 Qy 367 TATGAGGGGATGTTCTGACTCAAAAGTCCCTCATTTGAACACGTCGAAATCAGTGAACCA 426
 Db 370 TATGAGGGGATGTTTGAATGACGAAACCACTTGAACACGTCGAAATCAGTGAACCT 429
 Qy 427 TACTTGCGAAATCAAAATGATCCCAAGCTTTTCAAAAGGCGCATGACGCTGGTG 486
 Db 430 TATTGCGGATTAACGAATGATCCCAAGCTTTTCAAAAGGCGCATGACGCTGTA 489
 Qy 487 AAACGTTGGCTTGAATACGACAGCAACATTAAGAGTCAATTAAGAGCATGAG 546
 Db 490 AAACGAGTGGCTTGAATACGACAGCAAAATTAAGAGTCAATTAAGAGCATGAA 549

QY 547 GAAATGCTCAGTACGTAAGCAATGACGTCCATTATTAAGGCAAAAGCCTGAATAT 606
 DB 550 CAATATGCAATATTCGCAATAGCAATGATGTCTATATTAACGGCAAAAGCCTGAATAT 609
 QY 607 AAGGTATGAAATGATGTGGCCAGAGGTATTTGCAAAAGCGATGTGCAAGACAGCTAC 666
 DB 610 AAGGTATGAAATGATGTGGCCAGAGGTATTTGCAAAAGCGATGTGCAAGACAGCTAC 669
 QY 667 GGTGTATGACAAAGCCAGATGTTCGAGTTCCGATATCTGATTTGATACAGAGAA 726
 DB 670 GGTGTATGACAAAGCCAGATGTTCGAGTTCCGATATCTGATTTGATACAGAGAA 729
 QY 727 AACGACGTTCGATGATGATGAGCCTTCGGGTAAATTAACAGCACTATATGACAGTGGT 786
 DB 730 AATGACGTTCGATGATGATGAGCCTTCGGGTAAATTAACAGCACTATATGATGAGGAA 789
 QY 787 CGGACGATATGACGATGATGACGATGATGATGATGATGATGATGATGATGATGATGAT 846
 DB 790 CGGACGATATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
 QY 847 GGAATATGCGCAAGATTAAGGATGCACTCAAGCGAATCTGTTTTCATTCATC 906
 DB 850 GGAATATGCGCAAGATTAAGGATGCACTCAAGCGAATCTGTTTTCATTCATC 909
 QY 907 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
 DB 910 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 969
 QY 967 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
 DB 970 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
 QY 1027 TACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
 DB 1030 TACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
 QY 1087 TTCGGGCTGGGAAATGAAGGCGGAAAGCGGCGGATGATGATGATGATGATGATGATGATGATGAT 1146
 DB 1090 TTCGGGCTGGGAAATGAAGGCGGAAAGCGGCGGATGATGATGATGATGATGATGATGATGATGAT 1149
 QY 1147 AACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206
 DB 1150 AACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1209
 QY 1207 AATATTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
 DB 1210 AATATTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1269
 QY 1267 CCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
 DB 1270 CCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1329
 QY 1327 TCCCTCTGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386
 DB 1330 TCCCTCTGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1389
 QY 1387 AACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
 DB 1390 AACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1449
 QY 1447 AATCACTCTTAAGCCTTCTCTATTAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1506
 DB 1450 AATCACTCTTAAGCCTTCTCTATTAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1509
 QY 1507 TTGGGTTATCGGAAACGAAACGAAAGGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1566
 DB 1510 TTGGGTTATCGGAAACGAAACGAAAGGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1569
 QY 1567 GTTGCCTATGTAAGCAATCGATGCTTATCACTAGCCTTAAGCAATATACCTT 1626
 DB 1570 GTTGCCTATGTAAGCAATCGATGCTTATCACTAGCCTTAAGCAATATACCTT 1629
 QY 1627 ACTGCAAGGCGGCAAGCAATTTGAAATCTCTGATGATGATGATGATGATGATGATGATGATGAT 1686

DB 1630 ACTGCTACTGCGGCAAGCCTTGAATAATCTCTGATGATGATGATGATGATGATGATGATGATGAT 1689
 QY 1687 ACTACTGCTTCTGTAACCTTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1746
 DB 1690 ACACTGCTTCTGTAACCTTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1749
 QY 1747 ACACTGCTTCTGTAACCTTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1806
 DB 1750 ACACTGCTTCTGTAACCTTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1809
 QY 1807 AACGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1866
 DB 1810 AACGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1869
 QY 1867 GCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
 DB 1870 GCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1923
 RESULT 6
 CQ888285 1923 bp DNA linear PAT 19-OCT-2004
 LOCUS CQ888285
 DEFINITION Sequence 2 from Patent EP1466970.
 ACCESSION CQ888285
 VERSION CQ888285.1 GI:54304586
 KEYWORDS
 SOURCE Bacillus sp. KSM-KP43
 ORGANISM Bacillus sp. KSM-KP43
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE 1 Sato, T., Okuda, M., Koyama, S., Izawa, Y. and Kobayashi, T.
 Alkaline protease
 TITLE Patent: EP 1466970-A 2 13-OCT-2004;
 JOURNAL Kao Corporation (JP)
 FEATURES
 source location/Qualifiers
 1..1923
 /organism="Bacillus sp. KSM-KP43"
 /mol_type="unassigned DNA"
 /db_xref="taxon:109322"
 1..1923
 /note="unassigned protein product"
 /codon_start=1
 /protein_id="CAH6162.1"
 /db_xref="GI:54304587"
 /translation="MRKKKRVPLSVLSAAILSTVALSNPSAGARNPDLDFKGIOTT
 TDAGKPSKOGCGAALPLVSEBNVCLPGLQKGLFTVAPNMLHIIOPNGPILETKO
 QLETKAILEDITPDYATVETGDKVSKATSTIEHVSVEPLPIYRIDPOLPTGAS
 ELVKAVLADTKQKNEVOLRGEIQAFASIDVLYITAKPEYKVMNDVARGIVADV
 AOSYGLGQGOQIVAVADDTGLDTRNDSSMHEAPFGKLTALVALGRTNVANDTNGHT
 HVASVGLNGSTNKMAPQANLVFOSINDSGGLGGLPSNLTSOAYASAGARLHTN
 SMGAIVNATYDSDNNDVYRKNDMTLLPAAGNCGPAGGNTLSAGTAKNMTVATE
 NLKRPVGSYAANINHVAPSSRGPTDRIPVAPETPILSASSIAPPSFVANTH
 DSKATVKGSTHATPVAAGNVAQLKEHVNRKRGITPKPSLLPAALIAADIDGLTYPN
 GNQGWGRTVLTKSLNVAIVNBSSTLSQKATSTPATAGRKXISLWVSAPASTTA
 SVTLVNDLIDLVITAPNGTOYVGNDFSPYNDNMGRNVENVFINAPOSGYTIEVOA
 YNPVGPOTPSIAIVN"
 1..618
 sig_peptide
 mat_peptide
 ORIGIN
 Query Match 76.9%; Score 1477.2; DB 6; Length 1923;
 Best Local Similarity 85.7%; Pred. No. 0;
 Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
 QY 7 AAGAAAGAGGTGTTTATCTGTTTATGACGTGACGAGATCTGTGACATGTTGATTA 66
 DB 10 AAGAAAGAGGTGTTTATCTGTTTATGACGTGACGAGATCTGTGACATGTTGATTA 69
 QY 67 AACATCCCTGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 126
 DB 70 AATAATCCATGCGAGGTGCGAAGATTTTATGATGATGATGATGATGATGATGATGATGATGAT 129

Y:

```
/protein_id="CAH6186.1"
/db_xref="GI:54305177"
/translation="MRKKKVLIVLSAAIISTVLAISNPAGNRPDIPKGIOTT
TDKAPSKGGQTGMAALFVLESBNVLPKIQKELSTVPANIKLHIIQFNGPLSETKQ
BLKAVKALIDYIPDYAVIIVBYEGDVKSATSTIEHVSVEPLPYTRIDPOLFTTGAD
EIVKAVLDITKQKNKEVOLRGJBOIAQFASINDVLTIDPEPKYKWNIDARQIVADV
AOSSYLGGQCOQIVAVDGLDTPGRNDSMEHAFRGKITALPALRTNNANDNGHGT
HVASVILGNGSTNKGMAPQANI.VPOSINDSGGLGIPENITLTFQOAYSAQARIHTN
SWGAVNGAYTTDSRNVDVYAKNDMTILPAAGNRPNGGTSABGTAKMATVQATE
NIRPSFGTADNINHVADPSRFPKTDGNIKPDVNA.PGTFILSANSISLAPSSFWANH
DSKIVAMGTSMAPIVAGNVAKLRHFPVKNIGITPKSILAAIAGADADIGLYAPN
GNOGWRVTLDSKLNVAYVNBSSLSSTGKATYSFLATGKELKLSLWSDAPASTHA
SVTLVNDLDELVTIAPNGIQYVGNDFTSFYNDWMDGRNNVNFVINAPOSQTTTIVQA
YNVPIGPOTFFSLAIYN"
1..618
619..1920
/product="unnamed"
```

ORIGIN

```
Query Match      76.9%; Score 1477.2; DB 6; Length 1923;
Beet Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
```

```
QY 7 AAGAGAAAGTGTCTTTTATCTGTTTATCAGCTGCAGCGATTCTGCACTGTGGATTA 66
DB 10 AAGAAAAGGTCTTTTATCTGTTTATCAGCTGCAGCGATTGTGCACTGTGGATTA 69
QY 67 AACATCCCTCGGCTGGTGCAGCAAGACCTTTTGATCTGGATTTTAAAGAAATTCAMAC 126
DB 70 AGTATTCATCTGCAGTGTGTCAGGAATTTTGATCTGGATTTCAAAGAAATTCAGACA 129
QY 127 ACAACGATGTGAGGTCTTTCACAAAGCGACAAACGATGCGCGTGCATTTCTGGTG 186
DB 130 ACAACTGATGCTAAAGTTCCTCAAGCAGGCGCAGACTGATGCTGCTTTCTGGTG 189
QY 187 GAGTCTGAATAATGTGAACCTTCTTAAAGGATGTCTAAAGAACTTGAACATGACGCA 246
DB 190 GAATCTGAATAATGTGAACCTCCCAAGGTTTGAGAGAACTTGAACATGACCTCGCA 249
QY 247 AATAATTAATCTCATATGTCTCAATTCATATGSCCCCATTTTGAAGAAACAAACAGAG 306
DB 250 AATAATTAATCTCATATGTCTCAATTCATATGAGCCAAATTTAGAAACAAACAGCAG 309
QY 307 CTAGAGACAACTGAGCAAGAAATTCCTGACATACCTCGATTTATGCAATATGTCCAG 366
DB 310 CTGAAAAAAGCAGGCGCAAGATTCCTGACATACCTGATTAATGCTTACATTTGCGAG 369
QY 367 TATGAGGGGAGATGTCACTCAAAAGTCGCTCATTTGAACATGTGAATCAGTGGAGCA 426
DB 370 TATGAGGGGAGATGTTAAGTCAAGACACAGACCATTTGAGCAGTGAATTCGTGAGCTT 429
QY 427 TACTTCCGAAATACAAATAGATCCCACTTTTCACAAAGCGCATGACGCTGTG 486
DB 430 TATTGCGGATATACAGATATAGATCCCACTTTTCACAAAGGGGCAATCAAGCTTGA 489
QY 487 AAAAGCTGGCGCTTGAATTCAGAGCAAGACATTAAGAAAGTGCATTTAAGAGGCAATCGAG 546
DB 490 AAAAGAGGCGCTTGAATTCAGAGCAAGAAATTAAGAGGTGCAATTTAAGAGGCAATCGAA 549
QY 547 GAATTCGCTCAGTACGTAGCAAGCAATGACATTCATTAATTAACGCAAAAGCTGAATAT 606
DB 550 CAATTCGCAATTCGCAATAGCAATGATGTCTATATTAATTCGCAAAAGCTGAATAT 609
QY 607 AAGGTAGTAATGATGTGCGCAGAGATATTTCAAGCGGATGTGCAACAGACAGCTAC 666
DB 610 AAGGTAGTAATGATGTGCGGGAATTTCAAGCGGATGTGCTCAGAGCAGCTAC 669
QY 667 GGTGTGTATGACAAAGGCGCAATTTGTGCGAGTTGCCGATATCTGCAATTTGATACAGAGAA 726
DB 670 GGTGTGTATGACAAAGGCAAGATCTGAGCGGTGCGGATACAGGGCTTGTATCAAGGTCCG 729
QY 727 AACGACGTTGATGATGATGAGCTTCCGCGGTAAATTAACAGCACTATATGACATGAGT 786
```

```
DB 730 AATGACAGTGTGATGATGATGAGCTTCCGCGGGAATAATTAAGTATATATGCAATTTGGA 789
QY 787 CGAAGAAATTAATGCGAAATGATTCGAACGGTCAATGCTACCAATGTGCGAGTTCGGATTA 846
DB 790 CGAAGAAATTAATGCGAAATGATTCGAATGCTATGCTAGCGCATGTGCTGGCTCCGATTA 849
QY 847 GGAATTTGGGCGCAACGAATTAAGAAATGCGACCTCAAGCGAATCTGATTTTTCATATCATT 906
DB 850 GGAATTTGGGCTCCACTTAATTAAGAAATGCGGCTCAAGCGAATCTGATTTTTCATATC 909
QY 907 ATGATATGACAGTGTGCGGCTTGAAGCTTGCCTTCAATCTGCAACCTTATTCAGCCAA 966
DB 910 ATGATATGACAGTGTGCGGCTTGAAGCTTGCCTTCAATCTGCAACCTTATTCAGCCAA 969
QY 967 GCATTCAGTGCAGGTGCGCAAAATTCATACAACTCTTGGGCGGCGACCGGTGAATGGGCGC 1026
DB 970 GCATTCAGTGCAGGTGCGCAAAATTCATACAACTCTTGGGCGGCGACCGGTGAATGGGCGC 1029
QY 1027 TACACGACGATTTCCAGAAATGTGATGATGATATGAAAGAAATGATATGACGATTCCT 1086
DB 1030 TACACGACGATTTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
QY 1087 TTCGCGGCTGGGAATGAAAGCGCGGACCATCATGATGACACTTGTACGCTTAA 1146
DB 1090 TTCGCTGCGGGAATGAAAGCGCGGACCATCATGATGACACTTGTACGCTTAA 1149
QY 1147 AACCCATTAACATGTGCGGCAACCGAAACCTGCGCTCAAGCTTGGGTTCTATGCAAT 1206
DB 1150 AATGCAATTAACATGTGCGGCTTACGGAACCTCCCGCAAGCTTGGGTTCTATGCGGAC 1209
QY 1207 AATAATTAACGATGTCAGATTCCTTCCGCTGCGCGCAAAAGATGGGGAATCAAG 1266
DB 1210 AATAATTAACGATGTCAGATTCCTTCCGCTGCGCGCAAAAGATGGGGAATCAAG 1269
QY 1267 CCTGATGTCAGGCGCGGACCATCATTTATGCAAGATCTTCTTTCACCCGAT 1326
DB 1270 CCGATGTCATGCGCAACCGGAACGTTCAATCATGAGCAAGATCTTCTTTCACCCGAT 1329
QY 1327 TCTCTCTTTCGCGGCAATCATGACGCAATATGCTTATGAGTGAACGTCATGCA 1386
DB 1330 TCTCTCTTTCGCGGCAACCATGACGTAATATGATCATGATGAGTGAACGTCATGCT 1389
QY 1387 ACACGCAATTTGCGGGAATGTTGACAGCTCCGCGGCAATTTGTGAAAATAGAGCA 1446
DB 1390 ACACGCAATTTGCGGGAATGTTGACAGCTCCGCGGCAATTTGTGAAAATAGAGCA 1449
QY 1447 ATCACTCTTAAGCTTCCCTTATGGAACAGCTTGTGATTTGAGGTGCTGATGTTGA 1506
DB 1450 ATCACTCTTAAGCTTCCCTTATGGAACAGCTTGTGATTTGAGGTGCTGATGTTGA 1509
QY 1507 TTGGGTTATCCGAAACGAAACCAAGATGCGCGGCGAGTGAACCTGATTAATGTTGAAC 1566
DB 1510 CTGGCTACCGGAACGATTAACCAAGATGCGGGAAGAGTGAACCTGATTAATGTTGAAC 1569
QY 1567 GTTGCCTATGGAACGAATTCAGATGCTTATCAATACCAAAAGCGACATATACCTTT 1626
DB 1570 GTTGCCTATGGAACGAATTCAGATGCTTATCAATACCAAAAGCGACATATACCTTT 1629
QY 1627 ACTGCAACGCGGCGCAACGATTAAGAAATCTCCCTGATGATGATGATGATGATGATGATGATGAT 1686
DB 1630 ACTGCTATGCGCGGCAACGCTTGAATATCTCCCTGATGATGATGATGATGATGATGATGATGAT 1689
QY 1687 ACTACTGCTTCTGTAACCTGCTCAATGATTTGATTTGCTATTAACGACCAACGGA 1746
DB 1690 ACAACTGCTTCCGTAACCTTGTCAATGATCTGACCTTGTCAATACGCTCAATGAC 1749
QY 1747 ACAAGATATGTCGGAATGATCTTCAGACCAATTTGAACAATACTGGAGTGGCGGCAAT 1806
DB 1750 ACAAGATATGTCGGAATGATCTTCACTTCATCAATGATTAACCTGGAGTGGCGGCAAT 1809
QY 1807 AACGTAGAAATATGATTTATTAATTCGCGCAAAAGTGAACATATACATTTGAGGTGCA 1866
DB 1810 AACGTAGAAATATGATTTATTAATTCGCAACCAAGCGGGAAGTATCAATTTGAGGTACAG 1869
```

Y:

Query 1867 GCATTAATGTCGCCGTTGACACCAAACTTCTGTTGGCAATTGTGAACCTAA 1920
Db 1870 GCTTATTAACGTACCGGTTGACCAACGACCTTCTGTTGGCAATTGTGAATTTAA 1923

RESULT 8
AR368117
LOCUS AR368117 1923 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 5 from patent US 6376227.
ACCESSION AR368117
VERSION AR368117.1 GI:34601778
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 5 23-APR-2002;
Kao Corporation; Tokyo;
JPX;

FEATURES
SOURCE 1..1923
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 76.9%; Score 1477.2; DB 6; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Query 7 AAGAAAGAGTGTATTATCTGTTTATCACTGACGAGCGATTCTGTCATCTGTGCATTA 66
Db 10 AAGAAAAGTGTTTATCTGTTTATCACTGACGAGCGATTCTGTCATCTGTGCATTA 69
Query 67 AACATCCCTGGCGTGTGATGACGACGCTTTGATCTGTGATTTTAAAGAAATTCACAA 126
Db 70 AGTAATTCATCTGACGGTGTGACGAAATTTGATCTGTGATTTTAAAGAAATTCACAA 129
Query 127 AACACCGATGTCAGTGTCTTCCCAACGACGCAAAACAGGTGCGGCTGCAATTTCTGTG 186
Db 130 ACAACTGATGCTAAAGTTTCTCCACAGCGGGGACGACTGGTGTCTTCTGTG 189
Query 187 GAGTCTGAAAATGTGAACTTCTTAAAGATTTCTAAAGAACTTGAACAGTACCGGCA 246
Db 190 GAATCTGAAAATGTGAACTCCCAAAAGTTTGCAGAAAGCTTGAACAGTCCCGGCA 249
Query 247 AATAATTAACCTCCATATGTGCAATTCATGAGCCCATTTTAAAGAAACAAACAGAG 306
Db 250 AATAATTAACCTCCATATGTGCAATTCATGAGCCCATTTTAAAGAAACAAACAGAG 309
Query 307 CTAGAGCAACTGAGCAAAAGTTCTCGACTACATCCCTGATTTATGCAATATATTGTGAG 366
Db 310 CTGGAABAAAACAGGGGCAAAAGTTCTCGACTACATCCCTGATTTATGCAATATATTGTGAG 369
Query 367 TATGAGGGGAGTGTCTCAAAAGTCCGCTCCATTGAACAGTGAATCAGTGAAGCA 426
Db 370 TATGAGGGGAGTGTCTCAAAAGTCCGCTCCATTGAACAGTGAATCAGTGAAGCA 429
Query 427 TACTTGCAGAAATCAAAATAGATCCCGAGCTTTTCAAAAGGGGCAATCCGCTGTG 486
Db 430 TATTTCGCAATTAACAGATAGATCCCGAGCTTTTCAAAAGGGGCAATCCGCTGTG 489
Query 487 AAACGTTGGCGCTTGTATACAGAGCAAAATATAAGAGTGAATTAAGAGCATCGAG 546
Db 490 AAACGATGGGCGCTTGTATCAAAAGCAAAATATAAGAGTGAATTAAGAGCATCGAA 549
Query 547 GAAATCGCTCAGTACGTACAGCAAGATGACGTCAATTAATTAACGCAAGCTGAAATAT 606
Db 550 CAAATCGCAATTCGCAATTAAGCAATGATGTGTAATTAATTAACGCAAGCTGAAATAT 609

Query 607 AAGGTATGATATGATGTGGCCAGAGGTATTGTCAAGCGGATGTGGCAAGAGCACTAC 666
Db 610 AAGGTATGATATGATGTGGCCAGAGGTATTGTCAAGCGGATGTGGCAAGAGCACTAC 669
Query 667 GGTATTATGACAAAGCCAGATTTGTGCAAGTTCCGATCTGGAATTTGATACAGAGAGA 726
Db 670 GGTATTATGACAAAGCAAGATGTGAGCGGTTCCGATCTGAGGGCTTTGATACAGAGTGC 729
Query 727 AACGACAGTTGATGATGAAGCTTCCGCGGTAATAATACAGCACTATATGCACTGGGT 786
Db 730 AATGACAGTTGATGATGAAGCTTCCGCGGTAATAATACAGCACTATATGCACTGGGA 789
Query 787 CGAGCAATATGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
Db 790 CGAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
Query 847 GGAATGGCGGCAAGATTAAGAAATGAGCACTGCAAGCGATCTGTTTTTCAATCCATC 906
Db 850 GGAAGCGGCTCACTAATTAAGAAATGAGCACTGCAAGCGATCTGTTTTTCAATCCATC 909
Query 907 ATGATATGCAAGTGTGGCTTTGAGGCTTCCATCTGCAATCTGCAACCTTATTCAGCCA 966
Db 910 ATGATATGCGGTGGGAGCTTGAAGACTTCCGATCTGCAACCTTATTCAGGCCAA 969
Query 967 GCATTCAGTGCAGGTGCGAATTCATACAACTCCGCGGGGCGAGCGGTGAATGGGGCC 1026
Db 970 GCATTCAGTGTGTGCGAATTCATACAACTCCGCGGGGAGAGCAAGTGAATGGGGCT 1029
Query 1027 TACACGACAGATTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1086
Db 1030 TACACACAGATTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1089
Query 1087 TTGCGGCTGGAAATGAAGCGGACGCGGATCCATCAAGTCACTGTCAGCTGTCGCTTAA 1146
Db 1090 TTGCGCTCGGGAATGAAGCAAGCGGGAACCATCAAGTCACTGTCAGCTGTCGCTTAA 1149
Query 1147 AAGGCAATTAACAGTGGGCGCAACGAAACCTGCTCAAGCTTGGTTCTTATGAGAT 1206
Db 1150 AATGCAATTAACAGTGGGCGCAACGAAACCTCGCCCAAGCTTGGTTCTTATGAGAT 1209
Query 1207 AATAATTAACAGTGTGACAGTTCTCTCCGTTGAGCCCGCAAAAGATGGCGAATCAAG 1266
Db 1210 AATAATTAACAGTGTGAGCAAGTTCTCTTCAAGTGAACGCAAAAGATGAGCGATCAA 1269
Query 1267 CTTGATGATGAGCGGCAAGGACATCAATTTATCAGAGATCTTCTTGCACCGGAT 1326
Db 1270 CCGGATGATGAGCGGCAAGGACATCAATTTATCAGAGATCTTCTTGCACCGGAT 1329
Query 1327 TCCCTCTTGGGGGAATCAATGACAGCAATATGCTATATGGGTGGAACGTTCATGGCA 1386
Db 1330 TCCCTCTTGGGGGAATCAATGACAGCAATATGCTATATGGGTGGAACGTTCATGGCT 1389
Query 1387 ACACCGATTTGTGGGGGAATGTGCAAGCTCCGTGAGCAATTTTGAATAATAGAGA 1446
Db 1390 ACACCGATTTGTGGGGGAATGTGCAAGCTCCGTGAGCAATTTTGAATAATAGAGGC 1449
Query 1447 ATCACTCTTAACCTTCCCTATTAAGACAGCTTTGATTTGAGAGTGTCTGATGTGGA 1506
Db 1450 ATCACTCTTAACCTTCTATTAAGACAGCTTTGATTTGAGAGTGTCTGATGTGGA 1509
Query 1507 TTGGGTTATCGAAACGAAACCAAGATGGGGGCGAATGACCCGATTAATGTTGAC 1566
Db 1510 TTGGGTTATCGAAACGAAACCAAGATGGGGGCGAATGACCCGATTAATGTTGAC 1569
Query 1567 GTTGCCTATGTAACGAATCAAGTGCCTTATCACTAGCCCAAAAGGACATATACCTTT 1626
Db 1570 GTTGCCTATGTAACGAATCAAGTGCCTTATCACTAGCCCAAAAGGACATATACCTTT 1629
Query 1627 ACTGCAACGGGGGCAAGCAATGAATAATCTCCCTGATGTGTGATGATGATGATGATGATGAT 1686
Db 1630 ACTGCTACTCCGGCAAGCTTTGAATAATCTCCCTGATGTGTGATGATGATGATGATGATGAT 1689
Query 1687 ACTACTGCTTCTGTAAACCCCTGTCAATGATTTGATTTGATTTGATTTGATTTGATTTGAT 1746

Db	1690	ACAAATGCTTCGTACCGCTTGTCAATATCTGGACCCTTGCAATTACCGCTCCAAATGGC	1749
Oy	1747	ACAAGATATGTCGGGAATGACTCTTCAGCACCATTTTGACAAATACTGGATGCGGCNAI	1806
Db	1750	ACACAGTAGTAGGAAAATGACTTTACTTGCATACAAATGATTAACCTGGATGGCCGCAT	1809
Oy	1807	AACGTAGAATAATGTTATTATTATTCGCCCCAAAGTGAACAATATACATTGAGGTGCA	1866
Db	1810	AACGTAGAAAAATGTTATTATTATTAAGCACCAAAACCGGAGCGTATACATTGAGGTACG	1869
Oy	1867	GCAATTAATGTCGCCGTTGGACCAACAATACTTCGTGGCAATTTGTAACCTAA	1920
Db	1870	GCTTATAACGTACCGGTTGGACCAACAGACTTTCGTGGCAATTGTGAATTA	1923
RESULT 9			
AR368118			
LOCUS	AR368118	1923 bp	DNA linear PAT 12-SEP-2003
DEFINITION	Sequence from patent US 6376227.		
ACCESSION	AR368118		
VERSION	AR368118.1	GI:34601779	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1923)		
TITLE	Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,		
JOURNAL	Shikata,S. and Nomura,M.		
	Alkaline protease		
	Patent: US 6376227-A 7 23-APR-2002;		
	Kao Corporation; Tokyo;		
	jpx;		

Oy	1567	GTGGCTCATGTGACGAAATCCAGTGCCTTATCACTAGCCAAACAAAGCAATATACCTTT	1626
Db	1570	GTGGCGCATGTGAAAGAGTCCAGTCTCTATCAACAACAAAGGACAGTATCGTTT	1629
Oy	1627	ACTGGAACGGCGGGCAAGCCATTGAAATATCCCTCGTATATGTCGGATATGCCCTTGACAGC	1688
Db	1630	ACTGCTACTGCGCGGAGACCTTTGAAATATCTCCCGTATATGCTCTGAATGCCCTTCGAGC	1689
Oy	1687	ACTACTGCTCTGTATACCCGTGCAATGTTTGAATTTGATATACAGACCAAAACGSA	1746
Db	1690	ACAACTGCTTCCTCGTAAACGTTTGCAATGATCTGGAACCTTGTCATTAACCGCTCCAAATGAC	1749
Oy	1747	ACAAATATATGTCGGAGATGACTCTCAGACCACTTTGACAAATTAATCGAGATGACCGCAAT	1806
Db	1750	ACACAGTATGTAGAGAAATGACTTTATCTTGCAATACATATGATTAATGGAGATGGCCGCAAT	1809
Oy	1807	AAACGTAGAAATGATATTTATTAATTGCGCCCAAGTGAACATATATCAATTGAGGTGCAT	1866
Db	1810	AAACGTAGAAATGATATTTATTAATGACACCAACAAACGGGACGTATATCAATTGAAGTACAG	1869
Oy	1867	GCATATATATGTGCGCGGTGGACACCAAAATCTCTGTGGCAATTTGTGAATTA	1920
Db	1870	GCTATATACGTAACCGGTGGACACCAAGAACTCTGTGGCAATTTGTGAATTA	1923

RESULT 10

AR562465

LOCUS	1923 bp	DNA	linear	PAT 08-OCT-2004
DEFINITION	AR562465 Sequence 5 from patent US 6759228.			

ACCESSION	AR562465
VERSION	AR562465.1
	GT:53976506

KEYWORDS .

SOURCE	Unknown.
--------	----------

ORGANISM: Unknown: [Inc] aag4

REFERENCE 1 (bases 1 to 1,000)

AUTHORS Takaiwa, M., Okuda, M.

Shikata, S. and Nomura, M.
Alkaloid biosynthesis

Patent: US 6759228-A 5 06-JUL-2004:
JOURNAL
LIFE
Molecular Processes

Kao Corporation; Tokyo;

JPX;

FEATURES	LOCATION/QUALITIES	1923	SOURCE
1. 1923			

```
/organism="unknown"
```

```
/mol_type="genomic DNA"
```

ORIGIN

Query Match	76.98; Score 1477
-------------	-------------------

Best Local Similarity 85.7%; Pred. No. 0;

Matches 1641; conservative 0; mismatches 273; indels 0; gaps 0;

7 AAGAGAAGTGTTTTATCTGTTTATCAGCTGCAGCATTCGTGCACTGTCATTA 66

1. The first part of the document is a title page. It contains the title "THE HISTORY OF THE UNITED STATES OF AMERICA" and the author "BY JAMES MADISON".

Db	10	AAGAAAAAGGCGTTTAACTCGTTTAACTGACGTGAGGATTTGTGCACGTGGCGTTA	69
Qy	67	AACATCCCTGGGCTGTGATGCAGGACTTTGATTCGATTTTAAAGAAATTCAAACA	126
Db	70	AGTATTCATTCGACGGTGTGCAAGAAATTTGATTCGATTTCAAGGAATTCAGACA	129
Qy	127	ACAAACCATGTCATGCTGTTTCTCCAAACAGCGACAAACAGTGGCGCTGCATTTCTGTG	186
Db	130	ACAACTGATGTGTAAGGTTTCTCCAAAGAGGCGAGCTGTGTGCTGCTTTCTGTGTG	189
Qy	187	GAGTCTGAAATGTGAAACTTCTTAAGATGTGCTTAAGAAACTTGAAACAGTACCGCA	246
Db	190	GAATCTGAATATGTGAATCTCCAAAGGTTTGACAGAAAGCTTGAAACAGTCCCGCA	249
Qy	247	AATATAAACCTCCATTTGTTCATTCATTCAGTGGCCCATTTTAAAGAAACAAACAGAG	306
Db	250	AATATAAACCTCCATTTATTCCAATTCATTCAGCAACCAATTTTAAAGAAACAAACAGAG	309

QY	307	TTAGAGCAACTGGAGGCAAAAGTTTCTGACATACCCCGATTATGCAATATATGTCGAG	366
Db	310	CTGAAAAAACAGGGGCAAAAGTTTCTGACATACATCCGATTATGCTTACATTTGTCGAG	369
QY	367	TATGAGGGGGATGTTCAGTCAAAAGTCGCGTCCATTGAAACGTGGAAATCAGTGGACCA	426
Db	370	TATGAGGGCGAATGTTAAATCAGGCAACAAAGCACTTGAAGCGTGGAAATCCGTGGACCT	429
QY	427	TACTTGGCGAAATACAAAATAGATGCCAGCTTTTCAAAAAGCGCATCGACGTGGTG	486
Db	430	TATTTGCCGATATACAGATAGATCCCAAGCTTTTCAAAAAGGGCATCAAGACTTGTAT	489
QY	487	AAAGGTTTGGCGCTTGATACGAAAGCAACAATAAGAAAGTGCATTAAAGAGCATTCGAG	546
Db	490	AAAGGAGTGGCGCTTGATACAAAGCAAAAATAAAGAGTGCATTAAAGAGCATTCGAA	549
QY	547	GAATTCGCTCAGTACAGCAAGCAAGTACGCTCATTTATATACGGCAAAAGCTGAAATAT	606
Db	550	CAATTCGCAATTCGCAATAGCATATAGCATATGTGCTATATATTAACGGCAAAAGCTGAGTAT	609
QY	607	AAAGTATGAATGATGTGGCCAGAGATTTGTCAAAAGCGGATGTGGACAGAGCAAGCTAC	666
Db	610	AAAGTATGAATGATGTGGCGGTGGAAATTTGTCAAAAGCGGATGTGGCTCAGAGCAAGCTAC	669
QY	667	GCTTTGTATGGAACAAGGCCAGATTTGTCCGATGTGCCGATCTCGAATTTGGATACAGAAACA	726
Db	670	GGTGTGTATGGAACAAGGCAAGATCTTACGATCGTGTCCGATACAGAGGCTTGATACAGGTCCG	729
QY	727	AACGCAAGTTGGATGATGATGAAGCCTTCGGCGGTAAATAACAGCACTATATACACTGGGT	786
Db	730	ATATGCAATTTGATGATGATGAAGCCTTCGGCGGAAATAATCTGCACTTATATGCACTTGGGA	789
QY	787	CGAGCAATTAATGCGAATGATACGAAACGCTCATGCTAACCCATGTGGCAGGTTCCGTATTTA	846
Db	790	CGAGCAATTAATGCGAATGATACGAAATGCTCATGCTAACCCATGTGGCTCGGCTCCGATTTA	849
QY	847	GGAATATGGCGCAACGAATAAAGGAATGGCACTCAAGGGAATCTGGTTTTTCAATTCATC	906
Db	850	GGAATATGGCTCACGAATAAAGGAATGGCGCTCAGGGGAATCTATGCTTCCAAATCTATC	909
QY	907	ATGATATGACATGGGTGGGCTTGGAGGCTTGCTTCCTCAATCTGCAAAACCTTATTCAGCCAA	966
Db	910	ATGATATGACGTTGGGTGGGCTTGGAGGCTTGCATCTTGAAATCTGCAAACTTATATTCAGCCAA	969
QY	967	GCATTCAGTGCAGGTGTCAGCAATTCATCAAACTCTCTGGGGGCGAGCGGTGATGGGGCC	1021
Db	970	GCATTCAGTGTGGTGGCCAGAAATTCATCAAACTCTCTGGGGGCGAGCGTGAATGGGGCT	1022
QY	1027	TACACGACAGATTCAGAAATGTGGATGATATGTAAAGAAAATGATATGACGATTCCTT	1088
Db	1030	TACACACAGATTCAGAAATGTGGATGATATGTGCGCAAAAATGATATGACGATTCCTT	1089
QY	1087	TTCCCGGCTGGGAATGAAGGCCCAAGCGCGGTACATCAGTGCACCTGTACCGGCTAA	1144
Db	1090	TTCCGTGCTGGGAATGAAGGCAACGAAACGCGCGGAACATCAGTGCACAGGCAAGCTTAA	1144
QY	1147	AACGCCATTAACAGTTCGCGCAACCGAAACCTGCGTCAAGCTTTCGTTCTCATGCAAGAT	1201
Db	1150	AATGCAATTAACAGTTCGAGCTACGAAAACTCTCGGCCCAAGCTTTCGTTTATGCGAGC	1201
QY	1207	AATAATTAACAAGTTTGCAAGTTCCTTCCGTGGCCGCAAAAAGATGGCGGAATCAAG	1266
Db	1210	AATAATTAACAAGTTTGCAAGTTCCTTCAAGTTCGCAAGATTCGCAAGATTCGCAAG	1266
QY	1267	CCGATATGTCATGGGGCCAGGACATACATTTTATACGCAAGATCTTCTCTTGCAATCCGAT	1322
Db	1270	CCGATATGTCATGGGCAACCGGGAACGTTCTATATCAGCAAGATTCCTCTTGCAACCGGAT	1322
QY	1327	TCCGTCCTTCTGGGGGAATACGACAGCAAAATATGCTATATATGGGTGGAAACGTCCATGGCA	1388
Db	1330	TCCGTCCTTCTGGGGGAACCATGACAGTAAATATGATATCATATGGGTGGAAACGTCCATGGCT	1388
QY	1387	ACACCGAATGTTTGGCGGGAATGTTGCAACGCTCCGTGACATTTTGTGAAAAATAGAGGA	1444

• •

Db 1330 ACACCGATCGTTGCTGCGAAGCTGGCAACGTTTCGTAGCATTTTGTGAAAAACAGAGC 1449

Qy 1447 ATCACTCCCTAAGCCCTTCCCTATTTGAAGCAGCTTGATTTGCAGTGCTGTGATTTGGA 1506

Db 1450 ATCAACCAAGCTTCTCTATTAATAAGCGCAGTGAATGGCGGTGACGTGACATGGC 1509

Qy 1507 TTGGGTTATCCGAAACGAAACCAAGATGGGGCCGAGTGACCTTGATTAATCGTTGAAC 1566

Db 1510 CTGGGCTACCCGAAAGGTTAACCAAGATGGGGACAGTGAACATTTGATTAATCCCTGAAC 1567

Qy 1567 GTTGCGTTATGGAAGGAATCCAGTGGCCGATGAATAGCCAAAGGCAATATACCTTT 1622

Db 1570 GTTGCGTTATGGAAGGATCCAGTCCAGTCTCTATCCACAGCAAAAAGCAACGATCGTTT 1629

Qy 1627 ACTGCAACGGCGGGCAGCAATGGAATACTCCCTGATATGGTCGATAGCCCTGCAAGC 1688

Db 1630 ACTGCTATCGCGCGCAAGCTTGAATAATCTCCCTGATATGGTCGATAGCCCTGCAAGC 1689

Qy 1687 ACTATGCTCTCTGTAAACCTGGTCAATGATTTGGATTTGGTCATTAACGAACCAACGGA 1746

Db 1690 ACAATGCTCTCCGTAAACCTGTCAATGATCTGGAACCTTGTCATTAACCGCTCCAAATGGC 1749

Qy 1747 ACAAGATATGTCGGGAATGACTTCCACAGCAACATTTGACATTAATGGGAATGGCCGCAAT 1806

Db 1750 ACAAGATATGTAAGAAATGACTTAATCTTCCGCAATCAATGATTAATCTGGGAATGGCCGCAAT 1809

Qy 1807 AACGTAGAAAATGTAATTTATTAATTTGCGCCCAAGTGAACATATACATTTAGAGGTGCA 1866

Db 1810 AACGTAGAAAATGTAATTTATTAATTAATGACCAACAAAGCGGAGGTATACATTTAGAGTACAG 1867

Qy 1867 GCATTAATATGTCGGGTGAACCAACAAACTTCTGTTGGCAATTTGAACTAA 1920

Db 1870 GCTTAATACGTACCGGTGGAACCAACAGACTTCTGTTGGCAATTTGAACTAA 1923

QY	1267	CCGATGATGTCAGGCGCCGACGGGACATACATTTTATATGAGAAAGATCTCTCTTGACCCGAT	1328
Db	1270	CCGAGATGTCATGAGCACCGGAAAGCTTCACTATATGAGAAAGATCTCTCTTGACCCGAT	1329
QY	1327	TCCTCTCTTGCGGCGCATATCATGACAGCAAAATATGCTCTATATGAGTGAACGTCATGCGA	1386
Db	1330	TCCTCTCTTGCGGCGCAACCATATGACGTAAATATGCAATACATGAGTGAACGTCATGCT	1389
QY	1387	ACACCGAATGTTGTCGGGGGAATGTCACAGCTCCGTCAGCATTTTGTGAAAAATGAGGA	1446
Db	1390	ACACCGAATCGTTGCTGAGAAACGTGACACAGCTTCGTGAGCATTTTGTGAAAAACAGAGGC	1449
QY	1447	ATCAGCTCCGTAAGCCTTCCCTATATGAAAGAGGTTGATTTGACAGGTCGCTGATGTTGGA	1506
Db	1450	ATCAGACCAAAAGCCTTCTCTATATTAAGGGGACGTATTTGCGGTCAGCTGACATCGGC	1509
QY	1507	TTGGGTTATCCGAAACGAAACCAAGATAGGGGCGAGTCAGACCTTGATTAATTCGTTGAC	1566
Db	1510	CTTGGGTTATCCGAAACGTTATCCAAAGATAGGGGACAGATGACATTTGGATTAATCCCTGAC	1569
QY	1567	GTTGCGCTATGTGAAGGAATCCAGTCGCTTATTAATAGCCAAAAAGCGACATATACCTTT	1628
Db	1570	GTTGCGCTATGTGAAGAGTCGAGTCTCTATATCACAGCCAAAAAGCAGATACCTGTTT	1629
QY	1627	ACTGCAGCGGCGGGCAAGCATTTGAAAAATCTCCCTGTATGAGTGGAGTGGCCCTTGACAGC	1686
Db	1630	ACTGCTATCTGCGCGGCAAGCCTTTGAAAAATCTCCCTGTATGAGTGGAGTGGCCCTTGACAGC	1689
QY	1687	ACTACTGCTTCTGTAAACCCCTGTGCAATGATTTGATTTGTCATTACAGACCAAAACGGA	1746
Db	1690	ACAACTGCTTCCGTAAACGCTTGTCATGATCTGAGACCTTGTCATTACCGCTCCAAATGGC	1749
QY	1747	ACAAATATGTGGGAAATGACTTTCACAGACCATTTGACAAATATCGGGAATGGCCGCAAT	1806
Db	1750	ACACAGTATGTAGAAATGACTTTCCTGCAATGATTAATCTGGATGCGCGCAAT	1809
QY	1807	AACGTAGAAAAATGTAATTTATTAATTCGCCCCAAAGTGAACAATATACATTGAGGTGCA	1866
Db	1810	AACGTAGAAAAATGTAATTTATTTATATGACCAACCAAAAGCGGACGTATTAATTTGAACTACAG	1869
QY	1867	GCATATTAATGTGCGCGGTTGAGCACACAAAATTCTGTGGTGGCAATTTGTGAACATA	1920
Db	1870	GCTATTAATAGTACCGGTTGGAACACAGAACTTCTGTGGGCAATTTGTGAATTA	1923
RESULT 12			
LOCUS	AR069954	3003 bp	DNA linear PAT 18-FEB-2000
DEFINITION	Sequence 41 from patent US 5891701.		
ACCESSION	AR069954		
VERSION	AR069954.1	GI:7220842	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 3003)		
AUTHORS	Sloma,A. and Christensen,L.		
TITLE	Nucleotide encoding a polypeptide having protease activity		
JOURNAL	Patent: US 5891701-A 41 06-APR-1999;		
FEATURES	Location/Qualifiers		
source	1..3003		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match 59.2%; Score 1136.8; DB 6; Length 3003;			
Best Local Similarity 74.9%; Pred. No. 86-300;			
Matches 1336; Conservative 0; Mismatches 477; Indels 3; Gaps 1;			
QY	5	GAAAGATAGAGGTGTTTATCTGTTTATCAGCTGACAGCGATTCGTGCACTGTTGCAT	64
Db	859	GATCGAAGAGGATTTTTATCGTTTATCAGTGTGCACTATATGTTCTCTGTGCTT	918

QY	65	TAACAAATCCCTCGGCTGCTATGGAAGACATTTTGGATCTGGAATTTTAAAGAAATTCAA	124
Db	919	TAAAGCACTCCCTCTTAATAATGGGGCGAACATTTTGAATTGGAATTTTAAAGGATAGAA	978
QY	125	CAACAACCGATGTCAGTGGTTCTTCCAAACAGCGCAAAACAGTGCGGCTGCATTTCTGG	184
Db	979	CACCTTAGCCTAGAGAAAGGCTCGCACCAAGCAAGGAAAAACGGAAAGGCATCTTTTCTTG	1038
QY	135	TGGAGTCTGAAAATGTGAAACTTCTTAAAGGATTCCTTAAAGAACTTGAAACAGTACCGG	244
Db	1039	TAAACTCGAAAAATGTGAAAAATCCCAAGAAATATTCAAAAGAAACCTAGAACTAGTCCAG	1098
QY	245	CAAAATPAAACTCCATATTTGCTCAATTCATATGCGCCCATTTTGAAGAAACAAACAGA	304
Db	1099	CGGATPACAGACTATATATCGTTCAATTTGACGCACTATTTTGAAGGAAACGCAACTTC	1158
QY	305	AGCTTAGACACAATGAGACAAAGATTCTCGACTATCCCTGATTATGCAATATATTTGCG	364
Db	1159	AACTAGAAACAACGGGACGAAAAATCTCGATTATCTACCAAGATTACCTTATATTTGTCG	1218
QY	365	AGTATGAGAGGGGATGTTGACGCAAAAGTCCGCTCAATGAAACGATGGAATCAAGTGAAC	424
Db	1219	AATATGATGGGATGTAAAGCCGTACTTAAACGCAATTTGGCATTTTGGAAATCGGTTGAC	1278
QY	425	CATACCTTCCGAAATACAAAATAGATCCCAAGCTTTTCACAAAAGCGCATCGACGCTGG	484
Db	1279	CATATTTACCTTTATATAAATAAGACCCGCAATATTTTCAGAGAGACTTCTGAATTAG	1338
QY	445	TGAAAGCCTTGCGGCTGTATACGAAGCAGAACATTAAGAAATGTCATTAAGAGCAATCG	544
Db	1339	TAGAAACAGTAGCTTATAGATAAAAAGCAAAAGTAAAGAAATGCTTTAAGAGGATTTGG	1398
QY	545	AGCAATCGCTCAAGTACGTAGCAAGCATGACGTCCATATATTAACGCAAGCCTGAT	604
Db	1399	AACAAATTTGCCCAATACCGCACAAAATATGATTTATATAGTAAACCCAAAGCTGAT	1458
QY	605	ATPAGCTGATGATGATGTGCGCAGAGGTAATTTGCAAAACGGAATGTGCAACAGACAGCT	664
Db	1459	ACGAAGTTTTGAATGACGTGCGCCGCTGCATTTGTGAAGAACAGACGTGCACAAAATPACT	1518
QY	665	ACGGTTGTATAGCAAGAGCCAGATTTGTGCGAGTTGCGCACTGGAATTTGATACAGAA	724
Db	1519	TTGGCTTATATAGCAAGAGCAAGATTTGACAGTTGCTGATATCTGGGCTTTGATCAGAA	1578
QY	725	GAAACGACAGTTCGATGTCAGTAAGCCTTCGCGGTAAATTAACAGCACTATATGCACTGG	784
Db	1579	GAAATGACAGTTTCGATGTCAGTAAGCATTTCCGCGGTAAATTAACGCACTATATGCACTGG	1638
QY	785	GTCGACGCAATATATGCGAATGATACGAAACGGTCACTGTAACCATGTGCAAGTTCCGAT	844
Db	1639	GCAAGAACAAATPACCCCAATGATTCAAATGACATGAAACCCATGTTGCTGAAATCTGGT	1698
QY	845	TAGAAATAGCGCGCAACGAATTAAGGAATGACACTCAAGCAATCTGCTTTTCAATCCA	904
Db	1699	TAGAAAA---GCTAACAAATTAAGGAATGACACCGCAAGCCAACTAGTCTTTCAATCTA	1758
QY	905	TCATGATAGACAGTGTGAGGCTTGGAGGCTTGCTTCCATCTGCAACCTTATTCAGCC	964
Db	1756	TTATGATAGTGTGTGAGAGGGCTGGGAGGACTACCTGCTAACTTCAAAACATTAATTCAGTC	1815
QY	965	AAGCAATTCAGTGCAGGTGCGCAAGATTTATACAAACTCCCTGGGGGGGACGCGGTGATGGGG	1024
Db	1816	AAGCAATATATGTCCTGAGCGAAGATTCATACGAATTCATGGGGGGCTCAAGTAAACGCTG	1875
QY	1025	CCTACACGACAGATTCAGAAATGTGATGACTATGTAAGGAAAAATGATATGACGATTC	1084
Db	1876	CCTATAGACACAGACTCTCGAATGTGATGATTTATGTGAAGAAAAATGATATGACGATTC	1935
QY	1085	TTTTTCGCGGCTGGGAATGAAAGCGCGGAACCGCGGTACATCAATGTCACCTGTATACGCTTA	1144
Db	1936	TTTTTTCGCGGCGGAATATAGGACCAAGGTACCGGTACATGTCACATGTCACACAGAACAGAA	1995

QY	1145	AAAAAGCCATTAACAGTCGGCGCAACCGAAACCTGCGTCCAAAGCTTGGCTTCTATGACG	1204
Db	1996	AAAAATGCCATTACAGTGTGGGGCAACCGAAACCTACAGTCCAGCTTCGAATCTTATGCGG	2055
QY	1205	ATATATATTAAACAAGTTCACAGTCTCTCCCGTGGCCGCAAAAGATGGCGCATCA	1264
Db	2056	ATATATATTAAACAATTTCTCTCAATTCCTTCAACAGGTCCTACATGAAATGACATTA	2115
QY	1265	AGCTGATGTCATGGCGCCAGGGACATACATTTATACGAAAGATCTTCTCTGCAACCG	1324
Db	2116	AGCCGGAAGTCATGGCAACAGGATAGTATATCTCTGCTAGATCATCATTAAGTCCAG	2175
QY	1325	ATTCTCCCTTCCTGGCGGAATATGACACCAATATAGCCTATATAGGTGGAAACGTCATG	1384
Db	2176	ATTCTCAATCTGGGCAACCACTGATATTAATATAGCCTTACATAGGTGTACTTAAAG	2235
QY	1385	CAACACCGATTTGTTGCGGGAAATGTTGACAGCTCCGTGAGCATTTTGTGAAAAATAG	1444
Db	2236	CTACTTCATTTGTATAGCAGTATATGTTGACATTAAGGAGCATTTTGTGAAAAATAG	2295
QY	1445	GAATCACTCTTAAGCTTTCCTTATGAAGCAGCTTGTATGACAGTCTGTGATGTTG	1504
Db	2296	GGGTATCTCTTAAGCTTTCCTTATGAAGCAGCTTGTATGACAGTCTGTGATGTTG	2355
QY	1505	GATTGGGTATATCCGAACGAAACCAAGATGGGCGAGTGAACCTGGATTAATGCTTGA	1564
Db	2356	GACTTGGCTTCCAAATGATGTAACCAAGATGGGGAAGATACCTTATGATTAATCCCTAA	2415
QY	1565	ACGTTGCTTATGTGAACGATCCAGTGCCTATCAACATGACCAAAAGCGCATATACCT	1624
Db	2416	ATGTGCGATTGTGATTAACGAGCCCTTATCAACAGTCAAAACCAACATATTCGT	2475
QY	1625	TTACTGCAACGGCGGCAAGCATTTGAAATCTCCCTGTATATGTGCGATGCCCTGCA	1684
Db	2476	TTAAGGCTCAAGCTGGTAAACCTTAAATAATCACTGTTTGGTCAGATGACCAAGTA	2535
QY	1685	GCACTACTGCTTCTGTAAACCTGTGTAATGATTTGGATTTGCTCATACAGACCAACG	1744
Db	2536	GCACGACCGGATCACTAACTTATGAAATGATTAAGACTTATGATCACTGCAACCAATG	2595
QY	1745	GAAACAAGTATGTGGGGAATGACTTCTAGACCACTTGAATAATATCTGGGATGGCCGCA	1804
Db	2596	GAACTTAATAGCTGGAATATGCTTACAGCACGATATGAATACATTTGGGATGGCAGAA	2655
QY	1805	ATTAACGTGAATAATATTTATTTAATTCGCCCCAAAGTGAACAATATACATTGAGGTGC	1864
Db	2656	ACAACGTGGAATAATGTGTTATTCATATGCTCTCTCAAGGGAACGTATACATCGAATGTC	2715
QY	1865	AAGCATATATATGTGCGGTTGGACCAAACTTCTCGTTGGCAATTTGTAACATPA	1920
Db	2716	AGGCTTACATATGATACGATTAAGTCGCGAAACCTTTCTTTAGCGATTGTACATTA	2771
RESULT 13			
BD062155			
LOCUS	BD062155	3003 bp	DNA linear
DEFINITION	Nucleic acid encoding a polypeptide having protease activity.		
ACCESSION	BD062155		
VERSION	BD062155.1	GI:22607760	
KEYWORDS	JP 2001514529-A/39.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 3003)		
TITLE	Sloma, A. and Christlanon, L.		
JOURNAL	Nucleic acid encoding a polypeptide having protease activity		
COMMENT	Patent: JP 2001514529-A 39 11-SEP-2001;		
	NOVO NORDISK BIOTECH INC		
	PN JP 2001514529-A/39		
	PD 11-SEP-2001		
	PF 09-JUN-1998 JP 1999503145		

FEATURES	Source	Location/Qualifiers
ORIGIN		1..3003 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
Query Match	59.2%; Score 1136.8; DB 6; Length 3003;	
Best Local Similarity	74.9%; Pred. No. 8e-300;	
Matches 1436;	Conservative 0; Mismatches 477; Indels 3; Gaps 1;	
PR	12-JUN-1997 US	08/873479
PI	ALAN SLOMA, LINNÉ	CHRISTIANSON
PC	C12N15/57, C12N15/75, C12N9/54, C12K14/00	
CC	Strandedness: Single;	
CC	Topology: Linear;	
FH	Key	Location/Qualifiers
5	GAAGAAGAGAGCTTTTATCTCTTTTATCAGTCGACGCAATCTGTGCACTGTTCAT	64
859	GATCGAAGAGAGGTTTTTTTATCCGTTTTTATCAGTGTGCACTATGTGCTTCTGTGCTT	918
65	TAAACAATCCCTCGCTGTGATGATCAAGAACTTTTATCTGATTTTAAAGAAATTCAA	124
919	TAAAGATCTCTTCTACTATTTGGGGCGAACAATTTGAAATTTGAATTGAACTTTAAGGGATAGAA	978
125	CAACAACCGATGTCAGTGGTTTCTCAACACGCAACAACAGTGCAGCTGATTTCTGG	184
979	CACCTTAGCGCTAGAAAGGCTGCCACCAAGCAAGAAAGAAAGGCAATCTTTCTGTG	1038
185	TGGAGTCTGAAATATGTGAACCTTTTAAAGATGCTTAAAGAACTTGAACAGTACCGG	244
1039	TAAACTCTGAAAATGTGAATCCCAAGATATTCAAAAGAAACTGAATGTTCCAG	1098
245	CAATTAATTAACCTCATATTGTCCAAATTCATATGCGCCCATTTTGAAGAAACAAACAGA	304
1099	CGATTAACAGCTATATATCGTTCAATTTTGAAGCACTATTTTGAAGAAACGCACTTC	1158
305	AGCTAGAGACAACTGAGCAAAAGATTCTGACTCATCTCCGTATTAATGCAATATATGTCG	364
1159	AACCTAGAAAGACGGAGCGAAATTTCTGATTAATACAGCAATTACGCTTATATTTGTCG	1218
365	AGTATGAGGGGGAATGTTCACTCAAAAGTCCGCTCATTGAACACAGTGAATCACTGAGC	424
1219	AATATGATGGGGAATGAAGGCGCTTAATCAACGCAATTTGGCATTTTGAATCGTTGAC	1278
425	CATCTTGGCGAATTAACAATATGATCCCAAGCTTTCAAAAAGGGGCAATCGACGCTG	484
1279	CATTTTACCTTTATTAATAATATACCCGCAATTAATTTTCCAGAGGAGCTTTGAAATTAG	1338
485	TGAAGCGTTGGCGCTTGATTCGAAGACAGAAACAATAAAGATGCAATTAAGAGCATCG	544
1339	TAGAAACAGTATGCTTTAGATTAAGACAAAGATGAAGATGACGTTTAAGAGGATTTGG	1398
545	AGGAATTCGCTCACTAGCTAGCAAGCAATGACGTTCATTAATTAACGCAAAAGCTGAAT	604
1399	AACAAATTGCGCAATACGCGCAAAATATATATGATTAATACGTAACCCCAAAGCTGAAT	1458
605	ATAAGTGAATGAATGATATGTGGCCAGAGATATTGTCAAAAGGGAATTTGGCACAGACGCT	664
1459	ACGAAGTTTTGAATGAAGTGTGGCCCGTGGCAATTTGAAAGAGAGAGTGGCACAAATTAAT	1518
665	ACGGTTTGTATGACAGGCGCAAGCTTGTGCGAGTTGCGATTAATGAAATTTGATACAGAA	724
1519	TTGGCTTAATATGACAGAGACAGATTTTGAAGTGTGATATCTGGGCTTGAATACAGAA	1578
725	GAAGACGACAGTTGCGATGAAAGCTTTCCGCGTAAATTAACGACTATATATGACTGG	784
1579	GAATATGACAGTTGCGATGAAAGCAATTTCCGCGTAAATTAACGCACTATATATGACTGG	1638
785	GTGCGAGCAATTAATGCAATGATACGAACGCTCATGTTACCCATTTGGCAGGTTCCGTAAT	844
1639	GCAGAAAGAAATTAACGCAATGATCCAAATGAGACATGGAACCCCATGTTTGTGAGATCTGTGT	1698

845 TAGGAATGGGCGAAGATTAAGGATGGACCTCAAGCGAATCGTTTTCATCA 904
1699 TAGGAAT---GCTTCAAAATTAAGGATGGACCGGACCCAACTTAGTCTTCAATCTA 1755
905 TCATGATAGCAGTGGTGGCTTGGAGGCTTGCCCTTCGAATCTGCAAACTTATTCAGC 964
1756 TTATGATATGTTGATGAGGGCTGGGAGGCTACCTGCTTAATCTACAAACATTAATCAATC 1815
965 AAGCATTACATGTCAGTGGTGGATTCATCAAACTCTCGGGGGGCGAGCGGTAATGGG 1024
1816 AAGCATATATGCTGAGAGGAGAAATTCATCAAAATTCATGAGGGGCTCCAGTAAACGGTG 1875
1025 CCTACACGACGATTCAGAAATGTGATGATATGTAAGAAATATGATAGAGATTC 1084
1876 CCTATACGACGACTCTCGAAATGTGATGATATGTAAGAAATATGATAGAGATTC 1935
1085 TTTTCGCGCTGGGATGAAAGCGGACCGGATACCATCATGTCACCTGCTACGCTA 1144
1936 TTTTGGCGCGGAAATGAGGAGCCAGTACGATGATCATGTCACAGGAAACAGCA 1995
1145 AAAACGCCATTAACGTCCGCGCAACCGAAACCTTGCTGCAAGCTTCGATCTTATGCG 1204
1996 AAAATGCCATTAACGTGGGCGCAACCGAAACCTTACGTCGAAGCTTCGATCTTATGCG 2055
1205 ATTAATATTAACGAGTTGCAAGTTCTCTCCGATGGCCGCAAAAGATGGCGAATCA 1264
2056 ATTAATATTAACGAGTTGCTCAATTCCTTCAAGAGTCTTACAGATGAGCAATTA 2115
1265 AGCCTGATGTCATGCGCGCGCAAGATATTTATCAGCAAGATCTTCTTTCGACCG 1324
2116 AGCGGAGGTGATGCAACAGGATGATTAATCTCTGCTGATCATCATTAAGTCCAG 2175
1335 ATTCTCTCTTGGCGCAATCATGACAGCAATATGCTTATATGCTGGAACGTCATG 1384
2176 ATTCTCTCTTGGCGCAACCATGATATGAAATATGCTTACATGAGTGGTACTTCTATG 2235
1385 CAACACCGATGTTGCGGGGAATGTTGCAAGCTCCGATGAGATTTTGAAGAAATAG 1444
2236 CTATCTCAATTTAGAGGTATATGTTGCAATTAAGGAGCAATTTTGAAGAAATAG 2295
1445 GAATCTCTTCAAGCTTCCCTTATGAAAGCAGCTTGTATGTCAGTGTGCTGATGTTG 1504
2296 GGGTAACTCTTAAGCTTCCCTTAAAGCTGCTTAAATGTCAGTGTGCTGATGTTG 2355
1505 GATTGGGTTATCGAAGCGAAACCAAGATGGGGCGAGTGAACCTGATTAATGCTTGA 1564
2356 GACTTGGCTTTCGAATGTTGTAACCAAGATGGGGGAAGTAACTTAAATATCCCTTA 2415
1565 AGTTGGCCATGTTGAACGAATCAGTGGCTTATCAACTAGCAAAAGGAGCATATACCT 1624
2416 ATGTCGATTTGTAAGTAAGCAAGCCCTTATCAACAGTCAAAAGGAGCATATACCT 2475
1625 TTACTGCAACGGCGGCGCAAGCATTTGAAATCTCCCTGATGTCGATGCGCCCTGCA 1684
2476 TTAGGCTCAAGCTGTAACCTTTAAATATCACTGTTTGGTCAAGTGAACAGGTA 2535
1685 GCATTAATGCTTCTGTAACCTGCTCAATGATTTGATTTGCTATTAACAGCAAAAG 1744
2536 GCACACGCGCATCACTTAATTTAGTAATGATTTAGACTTAATATCACTGCAACAAATG 2595
1745 GAACGAATGATGGGGAATGCTCTGACGACCAATTTGACAAATACGTTGGATGGCCCA 1804
2596 GAACGAATGATGGGGAATGCTCTTACGACCACTGATATTAACATTTGGATGGCAAA 2655
1805 ATTAAGTAAATATTTATTTATTTATTCGCCCCCAAGTGAACATATACATTTAGGTC 1864
2656 ACAAGTGAATATTTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2715
1865 AAGCATATATGTCGCGGTTGACCAAACTTCTCTGTTGCAATTTGTAATTA 1920
2716 AGGCTTAACATGATACAGTAAGTCCGCAAACTTTCTTTAGCGATTTGTAATTA 2771

RESULT 14
AX839476
LOCUS AX839476 1305 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 2 from Patent EP1347044.
ACCESSION AX839476
VERSION AX839476.1 GI:39922766
KEYWORDS
SOURCE
ORGANISM Bacillus sp. KSM-KP43
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 Okuda, M.K., Sato, T.K., Saito, K.K., Sumitomo, N.K., Izawa, Y.K.,
Saei, K.K., Kobayashi, T.K. and Nomura, M.K.
TITLE Alkaline protease
JOURNAL Patent: EP 1347044-A 2 24-SEP-2003;
Kao Corporation (JP)
FEATURES
source
1..1305
/organism="Bacillus sp. KSM-KP43"
/mol_type="unassigned DNA"
/db_xref="taxon:109322"
<1..1305
/note="unassigned protein product"
/codon_start=1
/transl_table=1
/protein_id="CA85146.1"
/db_xref="GI:39922766"
/translation="NDVARGIVKADVAQSYGLYGQOIVAADTGLDTRGDSMEH
AFRQKIALYALGRITNNANDNGHGVASVLNGSTNKGMAQANLVFQSIDSG
GLGGLPSNLQTLFQAYSAGARLHNTNNGAIVNAYTTDSHYVDYKIDMTLLFA
GNEBNGGTTISAPETKATITVAGTENVLPFGSYANINHVAGESRGPTKDRKP
DVMAKPTTISABSSILA PDSFMAHNSKYAVMGSTMAPTPIVAGVAAALREHVKR
GITPKSLKALITAGAADTGLGTPNQGKRTVTLKSLNVAIVYSSSSISTQKAT
YSPFATGKPLKISLWSDAPASTTASVTTLVNDLVIITANGTQYVGNDFSTPYN
MDGNVNVNFINAPSGTYTIEVQANVVPQPTFSIAIVN"
ORIGIN
Query Match 52.2%; Score 1002.6; DB 6; Length 1305;
Best Local Similarity 85.5%; Pred. No. 4.8e-263;
Matches 1116; Conservative 0; Mismatches 189; Indels 0; Gaps 0;
QY 616 AATGATGGCGCAGAGATTTGTCAAAAGCGATGTGCAAGACGTAAGCTTTGTAT 675
DB 1 AATGATGGCGCGGTGAATTTGTCAAAAGCGATGTGCTCAGACGTAAGCTTTGTAT 60
QY 676 GAACAAGCGCAGATTTGTGCAAGTTGCCGATATCTGATTTGATACAGAAAGACAGT 735
DB 61 GAACAAGGACAGATCGTGTGCGGTTGCCGATATCAAGGCGTTGATACAGGTCGCAATGACAGT 120
QY 736 TCGATGCAATGAAGCTTCCGCGGTAAATTAACACATATATGCACTGGGTGCGACGAT 795
DB 121 TCGATGCAATGAAGCTTCCGCGGTAAATTAACATTTATATGCAATTTGGACGACGAT 180
QY 796 AATGCAATGATTAAGAACGCTGATGTACCATGTGGCAGGTTGCGATTAAGAAATGC 855
DB 181 AATGCCAATGATTAAGAACGCTGATGTACCATGTGGCTGCTCGATTTAGAAACGCG 240
QY 856 GCAACGAATTAAGAAATGCAACCTCAAGCGAATCTGTTTTCATTCATCATGATGATGC 915
DB 241 TCACTAATTAAGAAATGCGCGCTCAAGCGAATCTAATCTTCAATCATATGATGATGC 300
QY 916 AGTGTGGGCTTTGAGAGCTTCCCTTCAATTCGCAAACTTATTCAGGACGATTCAGT 975
DB 301 GGTGGGGGACTTTGAGAGCTTCCCTTCAATTCGCAAACTTATTCAGGACGATTCAGT 360
QY 976 GCAGGTGCGAATTCATACAACTCTGGGGGCGAGCGGTAATGGGGCTTACAGCA 1035
DB 361 GCTGTGCGAATTCATACAACTCTGGGGGCGAGCGGTAATGGGGCTTACAGCA 420
QY 1036 GATTTCGAATTTGATGATGATATGTAAGAAATATGATATGATGATGATTTCTTTGCGGCT 1095
DB 421 GATTTCGAATTTGATGATGATATGTAAGAAATATGATATGATGATGATTTCTTTGCGGCT 480

QY 1096 GGGAAATGAAGGCCGAAACCGCGGACATGACATGACATGCTGCTAAGGCTAAACCGCCATA 1155
 DB 481 GGGAAATGAAGGCCGAAACCGCGGACATGACATGACATGCTGCTAAGGCTAAACCGCCATA 540
 QY 1156 AAGATGCGGCGCAACCGAAACCTGCGCTCAAGCTTGGCTTCTATGACATATATATTAAC 1215
 DB 541 AAGATGCGGCGCAACCGAAACCTGCGCGCAAGCTTGGCTTCTATGACATATATTAAC 600
 QY 1216 CACGTTGCGACATGCTTCTCCGCGCGCGCAAAAGATGGCGCAATCAAGCTGATGTC 1275
 DB 601 CACGTTGCGACATGCTTCTCCGCGCGCGCAAAAGATGGCGCAATCAAGCTGATGTC 660
 QY 1276 ATGGCGCGCGGACATATATATATGAGAGATGCTTCTGACACCGATTCCTCTC 1335
 DB 661 ATGGCGCGCGGACATATATATGAGAGATGCTTCTGACACCGATTCCTCTC 720
 QY 1336 TGGGCGAATCATGACAGCAAAATATGCTATATGGTGGAACGTCATGCAACCGAT 1395
 DB 721 TGGGCGAATCATGACAGCAAAATATGCTATATGGTGGAACGTCATGCAACCGAT 780
 QY 1396 GTTGGCGGGAATGTTGCAACGCTCCGTGACATTTTGTGAAAATATGAGAAATCACTCT 1455
 DB 781 GTTGGCGGGAATGTTGCAACGCTCCGTGACATTTTGTGAAAATATGAGAAATCACTCA 840
 QY 1456 AAGCTTCCCTATGAAAGCAGCTTATGATGAGAGCTGCTGATGTTGATGCTTAT 1515
 DB 841 AAGCTTCCCTATGAAAGCAGCTTATGATGAGAGCTGCTGATGTTGATGCTTAT 900
 QY 1516 CGGAACGGAACCAAGATGAGGCGGAGTGAACCTGATATATGCTTGAACGTTGCTAT 1575
 DB 901 CGGAACGGAACCAAGATGAGGCGGAGTGAACCTGATATATGCTTGAACGTTGCTAT 960
 QY 1576 GTGAACGAATCCAGTCCCTATCAATGACCAAAAGCGACATATATCTTTATCTGCAAG 1635
 DB 961 GTGAACGAATCCAGTCCCTATCAATGACCAAAAGCGACATATATCTTTATCTGCTACT 1020
 QY 1636 GCGGCGAAGCCATGAAATCTCCCTGATATGTTGGAGTGGCCCTGCAAGCACTACTGT 1695
 DB 1021 GCGGCGAAGCCATGAAATCTCCCTGATATGTTGGAGTGGCCCTGCAAGCACTACTGT 1080
 QY 1696 TCTGTACCTGCTGATATGATTTGATTTGCTATTAACGACCAAAAGCGAAGATAT 1755
 DB 1081 TCTGTACCTGCTGATATGATTTGATTTGCTATTAACGCTTCAAAAGCGAAGATAT 1140
 QY 1756 GTGCGGAATGACTTCTGACGACCAATTTGACATTAACCTGGATGGCGGCAATATCGTAGA 1815
 DB 1141 GTGCGGAATGACTTCTGACGACCAATTTGACATTAACCTGGATGGCGGCAATATCGTAGA 1200
 QY 1816 AATGATTTATTAATTCGCCCAAGTGGAACATATATCATTTGAGGTGCAACATATAT 1875
 DB 1201 AATGATTTATTAATTCGCCCAAGTGGAACATATATCATTTGAGGTGCAACATATAT 1260
 QY 1876 GTGCGGTTGACACCAAACTTCTGTTGGCAATTTGAACTAA 1920
 DB 1261 GTGCGGTTGACACCAACCTTCTGTTGGCAATTTGAACTAA 1305

RESULT 15
 E03808
 LOCUS E03808 2218 bp DNA linear PAT 29-SEP-1997
 DEFINITION DNA encoding alkaline protease Ya.
 ACCESSION E03808
 VERSION E03808.1 GI:2172022
 KEYWORDS JP 1992197182-A/1.
 SOURCE Bacillus sp.
 ORGANISM Bacillus sp.
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE 1 (bases 1 to 2218)
 AUTHORS Tobe,S., Odera,M. and Arai,Y.
 TITLE DNA CODING ALKALINE PROTEASE YA ENZYME AND PRODUCTION OF ALKALINE
 PROTEASE YA USING THE DNA
 JOURNAL Patent: JP 1992197182-A 1 16-JUL-1992;
 LION CORP

COMMENT OS Bacillus sp.
 PN JP 1992197182-A/1
 PD 16-JUL-1992
 PF 28-NOV-1990 JP 1990327110
 PI TOBE SEIICHI, ODERA MOTOMASU, ARAI YOSHIO
 PC C12N15/57, C11D3/366, C12N9/54, (C12N15/57, C12R1/07), (C12N9/54, C12R1/07);
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: strain=Y strain;
 CC *source: clone=pUBA7;
 FH key Location/Qualifiers
 FH CDS 218..2125
 FT /product='precursor of alkaline protease Ya'
 FT 218..823
 FT mat_peptide 824..2122
 FT /product='Precursor of alkaline protease Ya'.
 FEATURES source Location/Qualifiers
 1..2218
 /organism='Bacillus sp.'
 /mol_type='genomic DNA'
 /db_xref='taxon:1409'
 ORIGIN
 Query Match 48.4%; Score 930; DB 6; Length 2218;
 Best Local Similarity 70.5%; Pred. No. 3,76-243;
 Matches 186; Conservative 0; Mismatches 530; Indels 9; Gaps 3;
 QY 96 TTTGATCGATTTTAAAGAAATCAACCAACCGATGCTAGTGTCTCCAAACA 155
 DB 310 TTTTCAAGTGAATTTTAAATGATGATAAAGTTTAAATAATG--CTAGTTGGTTAAACC 366
 QY 156 GCGAACAAGGTGGGCTGCACTTCTGTGAGTCTGAAAATGGAACCTCTTAAAG 215
 DB 367 GATTAAGCGGTGAGGCATCTTCTGTAGATATGGAATAATTAATATCTTAAAG 426
 QY 216 ATTGCTAAAGAACTTGAACAGTACCGGCAATATATTAATCATATTTGCTCAATTCA 275
 DB 427 TATTTAAAGAAAGCTAGAAAGCATGACAAAGATATACAACTTACATCTGATCAATTTAC 486
 QY 276 TGGCCCAATTTAGAAAGCAAAACAGAGCTAGAGCAACTGAGCAAGATTTCTGA 335
 DB 487 TGAACCAATTTCAAGAGAAAGCGAAAGATTAAGCTCTAGAGATATCATTTCTAGA 546
 QY 336 CTACATCCCTGATATATGATATATGTCAGATATGAGGGGATGTTCACTCAAAAGTCCG 395
 DB 547 TTATGTTCCAGATTAATGCTTTTATTTGTCAGAT--AGTGTGCTACAAAATAATTAAG 603
 QY 396 CTCCATGAAACGATGGAATCAGTGAAGCCATCTTGGCGAAATCAAAATATGATCCCA 455
 DB 604 TACTTTAATCTGTGTAAGAACGTACACATTTTACATTAATTAATTAATGATCTGA 663
 QY 456 GCTTTTCAAAAGCGCATGACGCTGTGAAACCGTTGGCGCTTGAATACAGACAGA 515
 DB 664 GCTTTTCAAAAGGTGCTTCCAGCTTGTTCAGCGGTTATTTAAATACAAAACAGA 723
 QY 516 CAATTAAGAAGTGCATTAAGAAGATGAGAAATGCTCAGTACGTAGCAAGCAATGA 575
 DB 724 AAATTAAGAAGTGAATTAATTCAGGTTTGAATGAGATGCTTCAATATGCTGCAAAATGA 783
 QY 576 GGTCCATATATTAAGCGCAAGCTGATATAGGTGATGATGATGTCGACAGAGTAT 635
 DB 784 TGTGTTTATATATACCAAAAGCCGAGTATGAGTATGATATGATATGCAAGAGGAT 843
 QY 636 TGTCAAGCGAGATGTGCAAGAGAGCTAGCGTTTGTATGACAAAGCCAGATTTGTCG 695
 DB 844 AGTAAGAAGCTAGTTGCAAAACATTAAGATTAATGAGATCAAGGTCAACTAGTTGC 903
 QY 696 AGTTGCCATCTGATTTGATACAGAAAGAAAGCAAGTTGATGATGAAGCTTCCG 755

Db 904 AGTAGCGGACACAGGCTTAGATACAGTGTCTAAGCATAGTCTATGATGAAAGCATTCGG 963
 Qy 756 CGGTAAATAAACAACATATATATGCACTGGTTCGACGAATTAATGCGAATGATACACGG 815
 Db 964 CGGGAATAATCACAGCTTTAGCGCTTAGAGAACTAATAATATGAGAGATCCGAATGG 1023
 Qy 816 TCATGTATCCCATATGCGGATGCTTCCGTATTTAGGAATGGCGCAAGATTAAGAAATGGC 875
 Db 1024 GCATGGCACACATGTAGCAGTTCGTACTTGTAT--GCTTTAAATTAAGAAATGGC 1080
 Qy 876 ACCTCAAGCGAATCTGTTTTTCAATCCATCATGATAGCAGTGTGGCTTGAAGCTT 935
 Db 1081 TCCGAAAGCTAATTAAGTCTTCCATCTAATATGATAGCAGCGAAGATTAAGTGGCTT 1140
 Qy 936 GCCTTCCAAATCTGCAACCTTATTCAGCCAGCATTCAGTGCAGGTGCCAATTCATAC 995
 Db 1141 ACCATCGAATTAATACGTTATTTATAGTCANAGCTTGAAATCTGGAGCAAGAAATCATAC 1200
 Qy 996 AAACCTCGGGGGGAGGCGGTGAATGGGGCCCTACACAGATTCAGAAATGTGATGA 1055
 Db 1201 TAACTGTTGGGAGCCCGCAATTAATGAGCGTACCTGTAACTGAGCAAGTGAATGA 1260
 Qy 1056 CTATGTAAAGAAATATATATGATACCATCTTTTCGCGCTGGGAATGAAGCCGAAACGG 1115
 Db 1261 GTATGTTGAAATATATATATGATAGCGTACTTTTGAGCTGGTAATGAAGTCTTAATTC 1320
 Qy 1116 CGGTACCATCAGTGCACCTGTGCGCTTAAACCGCATTAACATGCGCGCAACCGAAA 1175
 Db 1321 AGGAACAATATGATGTCTCCAGGTACAGCGAAATGTCTATACGTCGCGGCAACGAAA 1380
 Qy 1176 COTGCTCAAGCTTCGTTCTCTATGCAATTAATTAACAGTTGCAAGTTCCTTC 1235
 Db 1381 CTATGCCCAAGCTTCGTTCTCTATGCAATTAATTAACAGTTGCAAGTTCCTTC 1440
 Qy 1236 CCGTGGCCCGACAAATGAGGCGAATCAAGCTGATGTATGCGCGCAAGCAATACAT 1295
 Db 1441 GAGAGAGACTACGAGGAGATGAGCAATTAAGCTGACGTAAACGCTCTGGAAATTTAT 1500
 Qy 1296 TTTATCAGCAAGATCTTCTCTGCAACCGATTCCTCTTCTGGCGAATCATGACAGCA 1355
 Db 1501 TTTATCAGCAGTTCCTCTCTTAACTCCAGACTCTTCGTTTGGCGAATTAATTAACAGTA 1560
 Qy 1356 ATATGCTATATGGGTGAAAGTCCATGCAACAGATTTGTCGGGGAATGTGACA 1415
 Db 1561 ATACGCGTATATGAGCGGTAATCTCAATGCGACACCTATTTGTCAGGGAATGTCCGCA 1620
 Qy 1416 GCTCCGTAGCAATTTTGTGAAAAATAGAGAACTACTCTAAGCCTTCCCTATTGAAAGC 1475
 Db 1621 ATTACGTAGGCAATTTATTAATAAATAGAGTAATCTCTAAGCCTTCTTTAATAAAGC 1680
 Qy 1476 AGCTTTGATGAGGTGCTGCTGATGTTGGATTTCGGAACGGAACCAAGGATG 1535
 Db 1681 TGCACCTATCGTGTGCTACTGATGTGTTAGATATCTTAGGTGTGACCAAGGCTG 1740
 Qy 1536 GGGCCGAGTGAACCTGATTAATCGTTGAAAGTTGCTATGTGAAGCAATCCAGTCCCT 1595
 Db 1741 GGGGCGTGTACTCTAGATTAATCGTTAAATGTACGTAATGAGCAACTGCATT 1800
 Qy 1596 ATCAATGAGCCAAAAAGGAGCATATACCTTTACTGCAACGCGGGCAAGCCATTGAAT 1655
 Db 1801 AGCCACAGGACAAAAGCAACGTATTCGTTCCAGCACAAAGGGGTAAACCTTTAAAT 1860
 Qy 1656 CTCCCTGTATGTGGATGCCCCCTGCAAGCATATCTGTTCTGTAAACCTGTGTCAATGA 1715
 Db 1861 CTGCTATATGTGACAGATGTCTCGAAGTACAACTGCATCTTATACCTAGTTAATGA 1920
 Qy 1716 TTTGATTTGTCATTAAGCAACAAACGAAACAAGATATGTGGGAATGACTTCTCAGC 1775
 Db 1921 TTTAGATCTAGTATTTACTGCTCCGAAATGACAAAATATGTAGGAATGATTTTAGTTA 1980
 Qy 1776 ACCATTTGACATTAATCTGGATGCGCGCATTAACGTAGAAAAATGTATTTAATTCGCC 1835
 Db 1981 TCCATATATATTAATCTGGATGTGTGCAACATGTGGAAGCATTTATTAACGCTCC 2040

Qy 1836 CCAAGTGAACATATACCATTTGAGTGCAGACATATATATGTGCCGTTGGACCAAAA 1895
 Db 2041 GCAATCTGGAACGTATATTAATTTGAGTTCAAGCGGTATATATGTACCATCTGCCCAAGCG 2100
 Qy 1896 CTTCCTGTTGGCAATTTGAACTAA 1920
 Db 2101 TTCTCACTAGCTATGTACATTA 2125

Search completed: April 8, 2006, 06:48:23
 Job time : 9412.15 secs

activity over the range pH 6-12: (b) after 30 minutes at 40 deg. C it is

CC stable over the pH range 6-11; c) the isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
XX

Sequence 1920 BP; 576 A; 411 C; 473 G; 460 T; 0 U; 0 Other;

Query Match 100.0%; Score 1920; DB 2; Length 1920;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGAAAGAGAGAGAGTGTGTTTATCTGTTTATCAGCTGCAGCGATCTGTGACTGTT 60
DB 1 ATGAGAAAGAGAGAGAGTGTGTTTATCTGTTTATCAGCTGCAGCGATCTGTGACTGTT 60
QY 61 GCATTAAACATCCCTCGGCTGGTGTGATGCAAGCACTTTGATCTGGATTTTAAAGAAATT 120
DB 61 GCATTAAACATCCCTCGGCTGGTGTGATGCAAGCACTTTGATCTGGATTTTAAAGAAATT 120
QY 121 CAACAACACCGATGTCAGTGGTTCCTCAAAACAGCGACAAACAGTGGCGCTGCATTT 180
DB 121 CAACAACACCGATGTCAGTGGTTCCTCAAAACAGCGACAAACAGTGGCGCTGCATTT 180
QY 181 CTGGTGGAGCTGGAAGATGTGAACCTTCTTAAAGATTCCTTAAAGAACTTGAACAGTA 240
DB 181 CTGGTGGAGCTGGAAGATGTGAACCTTCTTAAAGATTCCTTAAAGAACTTGAACAGTA 240
QY 241 CCGGCAATTAATACTCCATTTGTCCATTCATTCATTCAGTGGCCCACTTTTGAAGAAACAAA 300
DB 241 CCGGCAATTAATACTCCATTTGTCCATTCATTCATTCAGTGGCCCACTTTTGAAGAAACAAA 300
QY 301 CAGAACTAGAGACAACTGAGCAAAAGATTCTGACTACATCCCTGATTATGCAATATTT 360
DB 301 CAGAACTAGAGACAACTGAGCAAAAGATTCTGACTACATCCCTGATTATGCAATATTT 360
QY 361 GTCGAGTATGAGGGGGAGTGTCAAGTCAAAAGTCCCTCATTTGAACAGTGGAAACAGT 420
DB 361 GTCGAGTATGAGGGGGAGTGTCAAGTCAAAAGTCCCTCATTTGAACAGTGGAAACAGT 420
QY 421 GAGCCATCTTCCGGAATTAACAATATAGATCCCGAGCTTTTCACAAAAGCGCATCGAG 480
DB 421 GAGCCATCTTCCGGAATTAACAATATAGATCCCGAGCTTTTCACAAAAGCGCATCGAG 480
QY 481 CTGGTGAAGCGTGGCGCTTGATTCGAAGCAAAATTAAGAAAGTCAATTAAGAGGC 540
DB 481 CTGGTGAAGCGTGGCGCTTGATTCGAAGCAAAATTAAGAAAGTCAATTAAGAGGC 540
QY 541 ATCGAGAAATCGCTCAGTACGTAGCAAGCAATGACCTTATNTATTAACGCAAAAGCT 600
DB 541 ATCGAGAAATCGCTCAGTACGTAGCAAGCAATGACCTTATNTATTAACGCAAAAGCT 600
QY 601 GAATATAAGTATGATGATGTGGCCAGAGTATTTCAAAAGCGAGTGGCAAGAGC 660
DB 601 GAATATAAGTATGATGATGTGGCCAGAGTATTTCAAAAGCGAGTGGCAAGAGC 660
QY 661 AGCTACGCTTTGTATGACAAAGCCAGTTGTGCGAGTTGCCGATATCTGAATTGGATACA 720
DB 661 AGCTACGCTTTGTATGACAAAGCCAGTTGTGCGAGTTGCCGATATCTGAATTGGATACA 720
QY 721 GGAAGAAACGACAGTGTGATGATGAGCTTCCGCGGTTAAATTAACGCACTATATGCA 780
DB 721 GGAAGAAACGACAGTGTGATGATGAGCTTCCGCGGTTAAATTAACGCACTATATGCA 780
QY 781 CTGGGTGCGAGCAATATGCAATGATGCAAGCGTCACTGTATCCCATGTGGCAGGTTCC 840
DB 781 CTGGGTGCGAGCAATATGCAATGATGCAAGCGTCACTGTATCCCATGTGGCAGGTTCC 840
QY 841 GTATTAGAAATGGCGCAAGATTAAGAAATGGCACTCAAGCGAATCTGCTTTTCA 900
DB 841 GTATTAGAAATGGCGCAAGATTAAGAAATGGCACTCAAGCGAATCTGCTTTTCA 900
```

```
DB 841 GTATTAGAAATGGCGCAAGATTAAGAAATGGCACTCAAGCGAATCTGCTTTTCA 900
QY 901 TCATCATGATATGACAGTGTGGGCTTGGAGGCTTGCTTCAATCTGCAAACTTATTC 960
DB 901 TCATCATGATATGACAGTGTGGGCTTGGAGGCTTGCTTCAATCTGCAAACTTATTC 960
QY 961 AGCCAGCACTTCAGTGCAGGAGGCCAAGTTCAATCAAACTCTGGGGGGGAGCGGTGAAT 1020
DB 961 AGCCAGCACTTCAGTGCAGGAGGCCAAGTTCAATCAAACTCTGGGGGGGAGCGGTGAAT 1020
QY 1021 GGGGCTTACACGACAGATTCAGAAATGTGATGATCTATGTAAGAAATATATATGACG 1080
DB 1021 GGGGCTTACACGACAGATTCAGAAATGTGATGATCTATGTAAGAAATATATATGACG 1080
QY 1081 ATTCTTTTCGGGCTGGGAATGAAAGCCGAACGGCGGTACATCAGTGCACCTGGTACG 1140
DB 1081 ATTCTTTTCGGGCTGGGAATGAAAGCCGAACGGCGGTACATCAGTGCACCTGGTACG 1140
QY 1141 GCTAAAAACGCGCATRACAGTGGGCGCAACGGAACCTGGGTCGAGCTTCGGTTCTAT 1200
DB 1141 GCTAAAAACGCGCATRACAGTGGGCGCAACGGAACCTGGGTCGAGCTTCGGTTCTAT 1200
QY 1201 GCAGATATATTAACAACGTTTGCACAGTTCTTCCGCTGGCCGACAAAAGATGGGCGA 1260
DB 1201 GCAGATATATTAACAACGTTTGCACAGTTCTTCCGCTGGCCGACAAAAGATGGGCGA 1260
QY 1261 ATCAAGCTGATATGTCATGCGGCCGAGGACATACATTTTATACGAAGATCTTCTTGCA 1320
DB 1261 ATCAAGCTGATATGTCATGCGGCCGAGGACATACATTTTATACGAAGATCTTCTTGCA 1320
QY 1321 CCCGATCTCTCCTTCTGGGCGAATCATGACAGCAATATGCTTATGAGTGAACGTCC 1380
DB 1321 CCCGATCTCTCCTTCTGGGCGAATCATGACAGCAATATGCTTATGAGTGAACGTCC 1380
QY 1381 ATGCGAACACCGAATGTTGGCGGGAATGTGCAAGCTCCGTGAGCATTTTGTGAAAAAT 1440
DB 1381 ATGCGAACACCGAATGTTGGCGGGAATGTGCAAGCTCCGTGAGCATTTTGTGAAAAAT 1440
QY 1441 AGAGAAATCACTCCTTAGCCCTTCCCTATGGAAGCACTTTGATGAGGTGCTGAT 1500
DB 1441 AGAGAAATCACTCCTTAGCCCTTCCCTATGGAAGCACTTTGATGAGGTGCTGAT 1500
QY 1501 GTTGATTTGGGTTATCCGAACGAAACCAAGATGGGGCCGAGTGCATCTGATTAATTCG 1560
DB 1501 GTTGATTTGGGTTATCCGAACGAAACCAAGATGGGGCCGAGTGCATCTGATTAATTCG 1560
QY 1561 TTGAACGTTGCTATGTGAACGAATCCAGTGCCCTATCAATGACCAAAAAGCGACATAT 1620
DB 1561 TTGAACGTTGCTATGTGAACGAATCCAGTGCCCTATCAATGACCAAAAAGCGACATAT 1620
QY 1621 ACCTTTATCTGCAACGGGGCGCAAGCCATGGAATCTCCCTGGTATGCTGGATGCCCT 1680
DB 1621 ACCTTTATCTGCAACGGGGCGCAAGCCATGGAATCTCCCTGGTATGCTGGATGCCCT 1680
QY 1681 GCAAGCACTACGCTCTGTATACCTTGTCATATGATTTGATTTGCAATTAACGACCA 1740
DB 1681 GCAAGCACTACGCTCTGTATACCTTGTCATATGATTTGATTTGCAATTAACGACCA 1740
QY 1741 AACGAAACAAAGATGTGCGGAATGACTTCTCAGCACATTTGAACATTAACGAGTGC 1800
DB 1741 AACGAAACAAAGATGTGCGGAATGACTTCTCAGCACATTTGAACATTAACGAGTGC 1800
QY 1801 CGCAATTAAGTAAAGATGTGCGGAATGACTTCTCAGCACATTTGAACATTAACGAGTGC 1860
DB 1801 CGCAATTAAGTAAAGATGTGCGGAATGACTTCTCAGCACATTTGAACATTAACGAGTGC 1860
QY 1861 GTGCAAGCATATATATGTGCGGTTGACCAAAACTTCTGTTGGCAATTTGCACTAA 1920
DB 1861 GTGCAAGCATATATATGTGCGGTTGACCAAAACTTCTGTTGGCAATTTGCACTAA 1920
```

RESULT 2
AAK37278

ID AAX37278 standard; DNA; 1923 BP.
XX
AC AAX37278: .
XX
DT 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
DE Bacillus alkaline protease encoding DNA.
XX
KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KM washing composition; oxidizing agent; ss.
XX
OS Bacillus sp.
XX
PN MO918218-rl.
XX
PD 15-APR-1999.
XX
PF 07-OCT-1998; 98WC-JP04528.
XX
PR 07-OCT-1997; 97JP-00274570.
XX
PA (KAOS) KAO CORP.
XX
PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
PI WPI, 1999-287736/27.
DR P-PSDB; AAY17088, AAY17090.
XX
XX
XX Alkali protease from Bacillus used in washing powders.
PS Disclosure, Page 58-63, 71pp; Japanese.
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
CC
XX
SQ Sequence 1923 BP, 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;

Query Match 76.9%; Score 1477.2; DB 2; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 7 AAGAAAGAGGTGTTTATCTGTTTATGAGCTGACGCGATTCGTGACTGTTCATTA 66
DB 10 AAGAAAGAGGTGTTTATCTGTTTATGAGCTGACGCGATTCGTGACTGTTCATTA 69
QY 67 AACATCCCTGGCTGGATGATGAGCAAGCACTTGTGATTCGATTTTAAAGGAATTCACA 126
DB 70 AGTAATTCATCTGACGAGTGATGAGCAAGCAATTTGATTCGATTTTCAAGGAATTCAGA 129
QY 127 ACAACCGATGTCAGTGTTCCTCAACAACGACGCAACAGTGCAGCTGCATTTCTGTG 186
DB 130 ACAATCGATGTCAGTGTTCCTCAACAACGAGGCGACGCTGCTGCTTTCTGTG 189
QY 187 GAGTCTGAAATATGAAATCTTTTAAAGGATTCGTAAGAAATTTGAAACGATCCGCA 246
DB 190 GAATCTGAAATATGAAATCTTCGAAAGGATTTGCAAGAAAGATTTGCAATCCCGCA 249
QY 247 AATAATTAACCTCATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 306
DB 247 AATAATTAACCTCATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 306

DB 250 AATTAATTAACCTCATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 309
QY 307 CTAGAGCACTGAGCAAGCAAGTTCGACATCATCTGATTTATGATTAATTCATTCATTCAT 366
DB 310 CTGAAAAAACAAGGCAAGATTCGACATCATCTGATTTATGATTAATTCATTCATTCAT 369
QY 367 TATGAGGGGGATGTTCAATCAAAAGTCCGCTCATTTGAACAGGTGAATCAGAGGACCA 426
DB 370 TATGAGGGGGATGTTCAATCAAAAGTCCGCTCATTTGAACAGGTGAATCAGAGGACCA 429
QY 427 TACTGCGCAATAACAAATATAGATCCGAGCTTTTCACAAAGGCGCATCGAGCTGTG 486
DB 430 TATTTGCGCATATACGAATATAGATCCGAGCTTTTCACAAAGGCGCATCGAGCTGTG 489
QY 487 AAAGCTTGGCGCTTGATTCGAAGCAAGCAATTAAGAAATGCAATTAAGAGCATCGAG 546
DB 490 AAAGCATGCGCTTGATTCGAAGCAAGCAATTAAGAAATGCAATTAAGAGCATCGAG 549
QY 547 GAAATCGCTCATAGTATGACGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 606
DB 550 CAAATCGCAATTCGCAATTAAGCAATGATTCGATTCGATTCGATTCGATTCGATTCGAT 609
QY 607 AAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
DB 610 AAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
QY 667 GGTGTTGATGACAAAGCCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 726
DB 670 GGTGTTGATGACAAAGCCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 729
QY 727 AACGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
DB 730 AATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
QY 787 CGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
DB 790 CGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
QY 847 GGAATGCGGCAAGCAATTAAGAAATGCACTCAAGCGAATCTGATTTTCAATTCATC 906
DB 850 GGAATGCGGCAAGCAATTAAGAAATGCACTCAAGCGAATCTGATTTTCAATTCATC 909
QY 907 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
DB 910 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 969
QY 967 GCATTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1026
DB 970 GCATTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1029
QY 1027 TACACGACGATTCGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
DB 1030 TACACGACGATTCGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
QY 1087 TTGCGGCTGGGAATGAAGCGCGCAACGCGGATCATCATGTCGCTTGATGAGCTTAA 1146
DB 1090 TTGCGGCTGGGAATGAAGCGCGCAACGCGGATCATCATGTCGCTTGATGAGCTTAA 1149
QY 1147 AACGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206
DB 1150 AACGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1209
QY 1207 AATATTAACAGTTCGACAGTTCCTTCGCGTGCACGCAAAAGATGAGCGAATTCAG 1266
DB 1210 AATATTAACAGTTCGACAGTTCCTTCGCGTGCACGCAAAAGATGAGCGAATTCAG 1269
QY 1267 CCTATGTCATGCGCGCCGCGGACATCATCTTTATGACGAAGTCTTCTTGACCCGAT 1326
DB 1270 CCGATGTCATGCGCGCCGCGGACATCATCTTTATGACGAAGTCTTCTTGACCCGAT 1329
QY 1327 TCTTCCTTCTGGGCGAATCATGACGAATATGCTTATGAGGATGAAAGTTCATGCA 1386
DB 1330 TCTTCCTTCTGGGCGAATCATGACGAATATGCTTATGAGGATGAAAGTTCATGCA 1389


```

QY 787 CGGACGAATATATGCGAATGATACGAACGCTATGCTATGCTGAGGTTCCGATTA 846
DB 790 CGGACGAATATATGCGAATGATACGAACGCTATGCTGAGGTTCCGATTA 849
QY 847 GGAAATGCGCGCAACGAATTAAGGAATGCACTCAAGCGAATCTGGTTTTCATATCATC 906
DB 850 GGAAACGGCTCCACTTAATTAAGGAATGCGGCTCAAGCGAATCTTACTTCCATATCATC 909
QY 907 ATGATATGAGTGTGTGGGCTTGGAGGCTTGCCTTCATCTGCAACCTTTATTCAGCCAA 966
DB 910 ATGATATGAGTGTGTGGGCTTGGAGGCTTGCCTTCATCTGCAACCTTTATTCAGCCAA 969
QY 967 GCATTCAAGTGAAGTGGCCGAATTCATACAACTCTGGGGGGGCGCGGTGAAATGGGGCC 1026
DB 970 GCATTCAAGTGTGTGGCCGAATTCATACAACTCTGGGGGGGCGCGGTGAAATGGGGCC 1029
QY 1027 TACACGACGATTCGAAATGTGTGATGCTATGTAAGGAATTAATGATGACATCTT 1086
DB 1030 TACACGACGATTCGAAATGTGTGATGCTATGTAAGGAATTAATGATGACATCTT 1089
QY 1087 TTCGCGCTGGGAATGAAGCGCGGAACGCGGTACCATCATGTCACCTGTGTAACGGCTAA 1146
DB 1090 TTCGCGCTGGGAATGAAGCGCGGAACGCGGTACCATCATGTCACCTGTGTAACGGCTAA 1149
QY 1147 AACGCGATTAAGTGGCGCAACCGAATCTGGCTCAAGCTTGGCTTCATGCAAGT 1206
DB 1150 AATGCAATTAAGTGGCGCAACCGAATCTGGCTCAAGCTTGGCTTCATGCAAGT 1209
QY 1207 AATATTAACGATGTGACAGTCTCTTCCGTGGCGCGCAAAAGATGGCGAATCAAG 1266
DB 1210 AATATTAACGATGTGACAGTCTCTTCCGTGGCGCGCAAAAGATGGCGAATCAAG 1269
QY 1267 CTTGATGTGATGGCGCGCAAGGACATATCATTTATTCAGCAAGTCTTCTTGACCCGAT 1326
DB 1270 CTTGATGTGATGGCGCGCAAGGACATATCATTTATTCAGCAAGTCTTCTTGACCCGAT 1329
QY 1327 TCTCTCTTCTGGGGGAATGATGACGAAATATGCTATATGCTGTAAGGAAACGTCATGGA 1386
DB 1330 TCTCTCTTCTGGGGGAATGATGACGAAATATGCTATATGCTGTAAGGAAACGTCATGGA 1389
QY 1387 AACACGATTTGTGCGGGGAATGTTGACAGCTCCGTGAGCATTTTGTGAAAATGAGAGA 1446
DB 1390 AACACGATTTGTGCGGGGAATGTTGACAGCTCCGTGAGCATTTTGTGAAAATGAGAGA 1449
QY 1447 ATCACTCTTAAGCTTCCCTATTTGAAGCACTTGTGATGAGTGTGAGTGTGGA 1506
DB 1450 ATCACTCTTAAGCTTCCCTATTTGAAGCACTTGTGATGAGTGTGAGTGTGGA 1509
QY 1507 TTGGGTTATCCGAACGGAACCAAGATGGGGCGGAGTGAACCTTGTGATTAATGCTTGAAC 1566
DB 1510 TTGGGTTATCCGAACGGAACCAAGATGGGGCGGAGTGAACCTTGTGATTAATGCTTGAAC 1569
QY 1567 GTTGCCTATGTAAGCAATTCAGTCCCTATCAATAGCCAAAGGCAATATACCTTT 1626
DB 1570 GTTGCCTATGTAAGCAATTCAGTCCCTATCAATAGCCAAAGGCAATATACCTTT 1629
QY 1627 ACTGCAAGCGCGGCAAGCATTTGAATCTCCCTGTGATGTGTGAGTGTGCAAGC 1686
DB 1630 ACTGCAAGCGCGGCAAGCATTTGAATCTCCCTGTGATGTGTGAGTGTGCAAGC 1689
QY 1687 ACTATGCTTCTGTGTAACCTGTGATGATTTGTGATTTGTGATTCAGCAACCAAGCA 1746
DB 1690 ACTATGCTTCTGTGTAACCTGTGATGATTTGTGATTTGTGATTCAGCAACCAAGCA 1749
QY 1747 ACAAGATATGTGCGGAATGACTTCTCAGCAACATTTGCAATTAATGAGGATGGCGCAAT 1806
DB 1750 ACAAGATATGTGCGGAATGACTTCTCAGCAACATTTGCAATTAATGAGGATGGCGCAAT 1809
QY 1807 AACGTAAGAAATGATTTATTAATTCGCGCCCAAGATGGAACATATACATTTGAGGTGCA 1866
DB 1810 AACGTAAGAAATGATTTATTAATTCACCAACGAGGAGCTATATCAATTTGAGTACAG 1869

```

```

QY 1867 GCATATTAATGTGCGGTTGGAACCAACAACTTCTGTGGCAATTTGAACTTA 1920
DB 1870 GCTTATTAACGTAACCGGTTGGAACCAACAACTTCTGTGGCAATTTGAACTTA 1923

RESULT 4
ADM40772
ID ADM40772 standard; DNA, 1923 BP.
XX
AC ADM40772;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding alkaline protease from Bacillus sp. KSM-KP43.
KW alkaline protease; laundry detergent; bleaching agent; detergent;
KW denure-cleaning agent; ds; gene.
OS Bacillus sp.; KSM-KP43.
XX
FH Key Location/Qualifiers
FT CDS 1..1923
FT /tag= b
FT /product= "Bacillus sp. KSM-KP43 alkaline protease"
FT sig_peptide 1..618
FT /tag= a
FT /product= "Alkaline protease prepro sequence"
FT mat_peptide 619..1923
FT /tag= c
FT /product= "Mature alkaline protease"
PN US2004072321-A1.
XX
PD 15-APR-2004.
XX
PF 09-JUN-2003; 2003US-00456479.
XX
PR 26-JUN-2002; 2002JP-00186387.
PR 18-OCT-2002; 2002JP-00304232.
XX
PA (KAOS ) KAO CORP.
PI Sato T, Okuda M, Takimura Y, Sumitomo N, Nomura M, Kobayashi T;
DR WPI, 2004-328572/30.
DR P-PSDB; ADM40773, ADM40770, ADM40771.
XX
PT New alkaline protease having a mutant prepro sequence where amino acid
PT residues at positions 52, 75 and 142 are substituted with another amino
PT acid residue, useful as enzyme component of laundry detergents, or
PT bleaching agents.
XX
PS Example 1; SEQ ID NO 3; 29pp; English.
XX
CC The invention relates to an alkaline protease having a prepro sequence.
CC The prepro sequence is a mutant sequence of SEQ ID NO: 1, or an amino
CC acid sequence having 80% homology or higher to SEQ ID NO: 1, where amino
CC acid residues at: (a) position 52 is substituted by aspartic acid or
CC arginine; (b) position 75 is substituted by alanine or arginine; and (c)
CC position 142 is substituted by lysine. The alkaline protease is useful as
CC an enzyme which can be incorporated into laundry detergents, bleaching
CC agents, detergents for cleaning hard surfaces or drapings, denure-
CC cleaning agents, and detergents for sterilising medical apparatus. The
CC present sequence represents DNA encoding alkaline protease from Bacillus
CC sp. KSM-KP43.
XX
SQ Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;

Query Match 76.9%; Score 1477.2; DB 12; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 7 AAGAAGAGAGGTGTTTATCTGTTTATCACTGCAAGGATTTGTGCACTGTGCATTA 66

```

[illegible]

Db		1090	TTCCGTGCGGGGAATGAAGAGCCGAACGGGCCAAACCATCACTGACCAAGGCACAGCTTAA	1149
Oy		1147	AACGCATTAACAGTCGGCGCAACCGAAAACTGCGTCCAGCTTCGGTTCCTATGCAGAT	1206
Db		1150	AATGCATTAAACATCGTGAGCTAACGGAAAACCTCGGCCAAGCTTTGGGTCTTAATGCGGAC	1209
Oy		1207	AATATTAAACAAGTTGACAGTTTCTCTCCGTGCCCCGACAAAAGATGGCGCAATCAG	1266
Db		1210	AATATTCAAACCAATGAGCACAGTTCTCTTCAAGTGAACGACAAAGATGACAGGATCAA	1269
Oy		1267	CCTGATGTCATGCGCGCCAGGAGCATPACTTTATCAGAAGATCTTCTCTTGACC CGAT	1326
Db		1270	CCGAGTATCATAGGACCGGGAAAGCTTCACTATCAGAGCAAGATCTTCTTTGACCGGAT	1329
Oy		1327	TCTCTCTTCTGCGCGAATCATGACAGCAAAATATSCCTATAATGSGTGNACGTCATGCA	1386
Db		1330	TCCTCTCTTCTGCGGGGAACCATGACAGTAATATGATCATATCATGSGTGGAAAGTCATGGCT	1389
Oy		1387	ACACCGAATGTTGCGGGGAATGTTGCACAGCTCCGTGACATTTTGTGAAAAATAGAGA	1446
Db		1390	ACACCGAATCGTTGCTGGAACAGTGGCACAGCTTCGTGAGCAATTTGTGAAAAACAGAGGC	1449
Oy		1447	ATCACTCTTAAGCCCTCCCTATTGAAAAGCAGCTTTGAATGAGAGTGCAGTATGTTGSA	1506
Db		1450	ATCAACACAAAGCCTTCTCTATTAAAGCGGCACTGATGTCGGGTGACGTGACATGGC	1509
Oy		1507	TTGGGTTATCCGAACCGGAACCAAGAGATGGGGCCGAGTGAACCTCGAATPAATCGTTGAC	1566
Db		1510	CTTGCGCTAACCGAACGGTAACGAAGATGGGGAAGATGACATTGGATPAATCCCTGAC	1569
Oy		1567	GTTGCGCTATGTTGAACGAATCCAAGTCCCCTATCAACTAGCCAAAAGGACATATACCTTT	1626
Db		1570	GTTGCGCTATGTTGAACGAATCCAAGTCCCCTATCAACTAGCCAAAAGGACAGTACTCGTTT	1629
Oy		1627	ACTGCAACCGGGGCAAGGACATGAAAAATCCTCGATATGTCGANTGCCCCGCAAGC	1686
Db		1630	ACTGCTATGCGCGCAAGCCTTTAATAATCTCCGTGATATGTCATATGCCCCGCAAGC	1689
Oy		1687	ACTACTGCTTCTGTAACCTGTGCAATGATTTGGATTTGGCTATTACAGACCAAAACGGA	1746
Db		1690	ACAATGCTTCCGTAACGGCTTGTCAATGATCTGGACCTTGTCAATTAACGCTCCAAATGGC	1749
Oy		1747	ACAAAGATATGTCCGGAATGACTTCTCAGACCAATTGACAAATTAATCGGAATGCGCAAT	1806
Db		1750	ACACAGTATGTAAGGAATGACTTTACTTCCTCGCATACATATATTACTGGATGCGCCGAAT	1809
Oy		1807	AACGTAGAAAAATGATTTATTAATTCGCCCCCAAGTGGACATATATACATTTGAGGTGCA	1866
Db		1810	AACGTAGAAAAATGATTTATTAATTAATGACCAACAAAGCGGACGTATACAAATTGATGATCAG	1869
Oy		1867	GCATATATATGTCGCGTTGACCACAAAACTTCTCGTTGGCAATTGGAATTA	1920
Db		1870	GCTTATTAACGTACCGGTGGACCAACACTTCTCGTTGGCAATTGGAATTA	1923
RESULT 5				
ID	ADSI4428			
	ADSI4428	standard; DNA; 1923 BP.		
XX	ADSI4428;			
AC				
XX				
DT	30-DEC-2004	(first entry)		
XX				
DE	Bacillus alkaline protease KP43 gene SPO ID NO:2.			
XX				
KM	protease; enzyme; alkaline protease; laundry detergent; KP43; ds; gene.			
OS	Bacillus sp.; KSM-KP43.			
XX				
FH	Key	Location/Qualifiers		
FT	FT CDS	1..1923		
FT		/carg= b		
FT		/product= "alkaline protease"		

Db 1510 CTGGCTACCGGAACGGTAACCAAGAGTGGGACGAGTGCATTTGGATTAATCCCTGAC 1569
QY 1567 GTTGCCTATGTGAACGAATCCAGTGCCTATCACTAGCCAAAAGCCAGATATACCTTT 1626
Db 1570 GTTGCCTATGTGAACGAATCCAGTGCCTATCACTAGCCAAAAGCCAGATATACCTTT 1629
QY 1627 ACTGCAAGGGGGGCAAGCATTGAAATCTCCCTGTATGGTGGGATGCCCTGCAAGC 1686
Db 1630 ACTGCTACTGCGGCAAGCCTTTGAAATCTCCCTGTATGGTGGGATGCCCTGCAAGC 1689
QY 1687 ACTACTGCTTCTGTAAACCTGTGTCAATGATTTGGATTTGATTAACAGCAACCAACGA 1746
Db 1690 ACAACTGCTTCCGTAAACGCTGTGTCAATGATTTGATTAACAGCAACCTTGTATACCGCTCAATGGC 1749
QY 1747 ACAAGATATGTGGGAATGATCTTTCAGACCAATTGACATTAACCTGGGATGGCCGCAAT 1806
Db 1750 ACACAGTATGTGAAGAAATGATCTTTCAGACCAATGATTAACCTGGGATGGCCGCAAT 1809
QY 1807 AACGTGAAGAAATGATTTATTAATTCGCCCCAAGTGAACATATACCTGAGGTGCA 1866
Db 1810 AACGTGAAGAAATGATTTATTAATGACCAACCAAGGGAAGTATACATTTAGGTAACG 1869
QY 1867 GCATTAATGTGCGCGTGGACCAAAAATTCTGTGGCAATTGTGAATTA 1920
Db 1870 GCTTAATGTGCGCGTGGACCAACAGACTTCTGTGGCAATTGTGAATTA 1923

RESULT 6

ADT49605
ID ADT49605 standard; DNA; 1923 BP.

AC ADT49605;

DT 30-DEC-2004 (first entry)

DE Bacillus alkaline protease encoding DNA.

KW Alkaline protease; detergent; fiber modification; leather processing;
KM cosmetic; bath additives; food-modification; pharmaceutical; gene; ds.

OS Bacillus sp. KSM-KP43.

Key location/Qualifiers

FT CDS 1..1923

FT /tag= b /product= "alkaline protease"

FT sig_peptide 1..618

FT /tag= a /note= "signal peptide"

FT CDS 619..1923

FT /tag= c /product= "mature alkaline protease"

FT /note= "SEQ ID 1"

EP1466970-A1.

PD 13-OCT-2004.

PP 08-APR-2004; 2004EP-00008605.

PR 10-APR-2003; 2003JP-00106708.

PA (KAO) KAO CO.®.

PI Sato T, Okuda M, Koyama S, Izawa Y, Kobayashi T;

DR MPI, 2004-711317/70.

DR P-PSDB; ADT49604, ADT49613.

XX New engineered alkaline protease with improved activity and thermal

PT stability, useful particularly in detergents such as laundry detergents.

XX

PS Example 2; SEQ ID NO 2; 35pp; English.

CC The invention relates to an alkaline protease and its encoding gene. The
CC alkaline protease can be expressed by standard recombinant methodology.
CC The alkaline protease is useful in the industry particularly in
CC detergents such as laundry detergents, but also in fiber modifying
CC agents, leather processing agents, cosmetic compositions, bath additives,
CC food-modifying agents, and pharmaceuticals. The enzyme has good activity
CC and thermal stability. The present sequence represents a DNA encoding an
CC alkaline protease from Bacillus sp. KSM-KP43.

XX Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;

QY Query Match 76.9%; Score 1477.2; DB 13; Length 1923;

XX Best local similarity 85.7%; Pred. No. 0; Mismatches 273; Indels 0; Gaps 0;

Db Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 7 AAGAAAGAGTGTGTTTATCTGTTTATCACTGACGAGATTTCTGTCATGTCATTA 66

Db 10 AAGAAAGAGTGTGTTTATCTGTTTATCACTGACGAGATTTCTGTCATGTCATTA 69

QY 67 AACATCCCTCGCTGTGATGCAAGACTTTGATCTGATTTTAAAGAAATTCACCA 126

Db 70 AGTATCCATCTGCAAGGTGTGCAAGAAATTTGATCTGATTTTCAAGAAATTCACCA 129

QY 127 AACACCGATGTCAGTGTGTTTCTCCAAACAGGACCAACAGGTGCGCTGATTTCTG 186

Db 130 ACAAAGTGTGTTTCTCCAAACAGGACCAACAGGTGCGCTGATTTCTG 189

QY 187 GAGTCTGAAATGTGAATCTTAAAGATTTGCTAAAGAACTTGAACAGTCCGCA 246

Db 190 GATCTGAAATGTGAATCTTAAAGATTTGCTAAAGAACTTGAACAGTCCGCA 249

QY 247 AATATTAATCTCATTTGTCTCAATTCATGAGCCCATTTTAAAGAAACAAACAGAG 306

Db 250 AATATTAATCTCATTTGTCTCAATTCATGAGCCCATTTTAAAGAAACAAACAGAG 309

QY 307 CTAGAGCAATCTGAGCAAGAAATCTGATCAATCTGATTAATGATATTTGTGAG 366

Db 310 CTAGAGCAATCTGAGCAAGAAATCTGATCAATCTGATTAATGATATTTGTGAG 369

QY 367 TATGAGGGGATGTTCACTCAAAAGTCCGTCATTAAGCAAGTGAACAGTGAAGCCA 426

Db 370 TATGAGGGGATGTTCACTCAAAAGTCCGTCATTAAGCAAGTGAACAGTGAAGCCA 429

QY 427 TACTTCCGAAATACAAATGATATCCAGCTTTTCAAAAGGCGCATCGAGTGTG 486

Db 430 TATTTGCCGATATACAAATGATATCCAGCTTTTCAAAAGGCGCATCGAGTGTG 489

QY 487 AAAGCTTGGCGCTGATACGAAGCAATTAAGAAAGTGAATTAAGAGCATCGAG 546

Db 490 AAAGCTTGGCGCTGATACGAAGCAATTAAGAAAGTGAATTAAGAGCATCGAG 549

QY 547 GAAATCGCTCAGTACGTAAGCAATGACGTCATTAATTAACGCAAGCTGAATAT 606

Db 550 GAAATCGCTCAGTACGTAAGCAATGACGTCATTAATTAACGCAAGCTGAATAT 609

QY 607 AAGGTGATGATGATGTGCGCAGAGTATTTCAAGGCGATGTGCAACAGCACTAC 666

Db 610 AAGGTGATGATGATGTGCGCAGAGTATTTCAAGGCGATGTGCAACAGCACTAC 669

QY 667 GGTGTTGATGCAAGGCGCAATTTGTCAGTGGCGGATATCTGATTTGATGATGAGAGA 726

Db 670 GGTGTTGATGCAAGGCGCAATTTGTCAGTGGCGGATATCTGATTTGATGATGAGAGA 729

QY 727 AACGACATTCGATGATGAAGCTTTCGCGGTAAATTAACAGCACTATATGATCGTGGT 786

Db 730 AACGACATTCGATGATGAAGCTTTCGCGGTAAATTAACAGCACTATATGATCGTGGT 789

QY 787 CGAGCAATTAATGCAATGATACGAAGCGTCAATGATCCATGTGCGAGTTCGATATTA 846

Db 790 CGAGCAATTAATGCAATGATACGAAGCGTCAATGATCCATGTGCGAGTTCGATATTA 849

[illegible][illegible]

Db 919 TAAAGATCTCTTCTATCTATTGGGCGAACAATTTGAATTGACCTTTAAGGGATAGAGA 978
Qy 125 CAACAACCGATGTCAGTGGTTTCTTCCAAACAGCGA CAACAGTGGCGCTGCATTTCTGG 184
Db 979 CACTTACCTTAGAAGAGCTGCCAAGCAAGAAAAACGGAAAGGATCTTTCTTCTGG 1038
Qy 185 TGGAGTCTGAAATGTGAACTTTAAAGATGTCTAAAGAACTTGAACAGTACCGG 244
Db 1039 TAAACTCGAAATGTGAAATCCCAAGATTTCAAAAGAACTAGAAAGTATTCAG 1098
Qy 245 CAATATTAATCTCATATTGTCATATTCATATTCGCGCCCATTTTGAAGAAACAAACAGA 304
Db 1089 CGGATTAACAGCTATATATCTGTTCAATTTGACGCGACTATTTTAAAGAAACGCACTTC 1158
Qy 305 AGCTAGAGACAACCTGAGCAAAAGATTCTGACTACATCCCTGATTTGCTATATTTCTG 364
Db 1159 AACTAGAGAGACGGGAGCGAAATTTCTGATTAATACAGATTACGCTTATATTTCTG 1218
Qy 365 AGTATGAGGGGATTTCAATCAAAAGTCCGCTCAATTGAACAGTGGATCAGTGGAC 424
Db 1219 AATATGATGGGATGTAAAGGCGTAACTAAACGCAATTCGCAATTTGGAATCGTGGAC 1278
Qy 425 CATATCTGCGAAATATACAAATATGATCCGAGCTTTTCAAAAAGCGCATGACGCTGG 484
Db 1279 CATATTTACTTTATATTAATATAGACCGGCAATTTATTTCCAGAGAGCTTCTGAATTAG 1338
Qy 485 TGAAGCGTTGGCGCTTGATACGAAGCAAAACAATTAAGATGCAATTAAGAGGATCG 544
Db 1339 TAGAAAACGATGCTTATGATTAAGAAAGCAAGAAAGTAAAGATGATTTAAGAGATTTGG 1398
Qy 545 AGGAAATGCTCAGTACGTAGCAAGCAATGACGTCAATTTATTTACGGCAAAAGCTGAT 604
Db 1399 AACAAATGGCCCAATACCGCAAAATATGATGATTTATACGTAAACCCCAAGCTGAT 1458
Qy 605 ATTAAGTATGATGATGTGGCCAGAGGATTTGTCAAAAGCGGATGGGACAGACAGCT 664
Db 1459 ACGAAGTTTGAATACGTGGCCGCTGGCATTTGTAAAGCAGACGTCCGCAAAATACCT 1518
Qy 665 ACGGTTGTATGACAAAGCCAGATTGTCCGAGTTGCCGATACCTGATTTGATACAGAA 724
Db 1519 TTGGCTTATATGACAAAGACGATTTAGCAGTTGTGATCTGAGCTGATACAGAA 1578
Qy 725 GAAACGAGATTCGATGATGAAAGCTTCCGCGTAAATTAACAGCACTATATGACCTGG 784
Db 1579 GAAATGACAGTTCCAGTATGAAAGCATTCGCGGTAAGATTAACGCACTATATGACCTGG 1638
Qy 785 GTCCGACGATATATGCAATGATACGAACGCTCATGATACCAATGTCAGGTTCTGAT 844
Db 1639 GCAAGACGATTAACGCCAATGATCCAAATGACATGGAACCAATGTCGATCTGTCT 1698
Qy 845 TAGAATATGGCGCAACGAATTAAGGATGGCACTCAAGCAATCTGTTTTTCATCCA 904
Db 1699 TAGAATAT---GCTACAAATTAAGGATGGCACCGCAAGCAATCTAGCTTTCAATCTA 1755
Qy 905 TCATGATATGCAATGTGTGGCTTGGAGGCTTGCCTTCCAAATCTGAAACCTTATTCAGCC 964
Db 1756 TTATGATATGATGTGTGAGGCTGGGAGGACTACCTGCTAATCTTAACCAATTTATTCAGTTC 1815
Qy 965 AAGCATTCAGTCAAGTGCAGAAATTCATACAAATCTCTGGGGGGCGACGGTGAATGGGG 1024
Db 1816 AAGCATATATGTCTGAGACGAAATTCATACGAATTCATGGGGGGCTCAGTAAACGCTG 1875
Qy 1025 CCTACACGACAGATTCAGAAATGTGATGACTATGTAAGAAAAATGATATGACGATTC 1084
Db 1876 CCTATACGACAGACTCTCGAATGTGATGATTAATGTAAGAAAAATGATATGACGATTC 1935
Qy 1085 TTTTTCGCGCTGGGATTAAGGCGGAAACGCGGTGATCAATCAGTGCACCTGTAGCGGTA 1144
Db 1936 TTTTTCGCGCGGAAATGAGGACCGAGTACGATCAATCAGTGCACGAGAAACAGGAA 1995
Qy 1145 AAAACGCAATACAGTGGCGCAACGAAACCTGCTCAAGCTTCTGATTCATGACG 1204

Db 1996 AAAATGCATTAACAGTTGGGCGAACCGAAACCTACGTCCAGCTTCGATCTTATGGG 2055
Qy 1205 ATTAATATTAACACGTTGGACAGTCTTCTCCGTGGCCGACAAAGATGGCGAATCA 1264
Db 2056 ATTAATATTAACAGTTGGCTCAATTTCTTCAAGAGCTCTACAGATGAGATTA 2115
Qy 1265 AGCGTATGCAATGCGCGCAGGGAATTAATTTATGAGAAAGTCTTCTGTCACCG 1324
Db 2116 AGCGGACGTATGCAACAGGTACGTATATTTCTCTGCTAGATCAATTAAGCTCAG 2175
Qy 1325 ATTCCTCTTCTGGGCGAATCATGACAGAAATATGCTATATGAGTGAACGTCATG 1384
Db 2176 ATTCCTCATTTCTGGGCAACATGATATATATGCTTACATGAGTGTACTTCTAG 2235
Qy 1385 CAACACGATTTGTTGGGGAATGTGACAGCTCCGTGACATTTTGTGAAAAATAGAG 1444
Db 2236 CTACTCCAAATTTGAGCAGGTATGTGACAAATTAAGGAGCAATTTGTGAAAAATAGAG 2295
Qy 1445 GAATCACTCTTAACCTTCCCTATTTGAAGAGCTTTGATTTGCAAGTCTGTGATTTG 1504
Db 2296 GGGTAACTCTTAACCTTCTCTTTTAAAGCTGCTTAATTTGCAAGTGTGCGAGTGTG 2355
Qy 1505 GATTGGTTATCCGAACGAAACCAAGGATGGGCGAGTGAACCTGGATTAATGTTGA 1564
Db 2356 GACTTGGCTTCCAAATGTATACCAAGATGGGGAAGATGATGATATTAATCCCTAA 2415
Qy 1565 AGTTGCTTATGTGACGAATCCAGTGCCTTATCACTAGCCAAAGCGACATATACCT 1624
Db 2416 ATGTGCAATTTGTGAATGAACGAGCCTTATCAACAGTCAAAAGCAATATTCGT 2475
Qy 1625 TTATGCAACGGCGGCGCAAGCATTTGAAATCTCCGTGATATGCTGGATGCCCTGCA 1684
Db 2476 TTACGGCTCAAGCTGTGTAACCTTTAAATTAATCACTGTGTGTCAGATACCAAGTA 2535
Qy 1685 GCATTAAGCTTCTGTAACCTTGGTCAATGATTTGATTTGATTTGATTAACAGCAACGA 1744
Db 2536 GCAAGAGGCAATCACTAATTTAGTATGATTTAAGCTTATGATCACTGACCAATAG 2595
Qy 1745 GAAACAGATATGTGGGAATGACTTCTCAGACCAATTTGACATTAATCTGGAGTGGCGCA 1804
Db 2596 GAACCTAATATAGTGGAAATGACTTTACAGCACGATATGATTAATTTGGGATGGCGAA 2655
Qy 1805 ATTAAGTAAAGAAATGATTTATTAATTCGCCCGAAAGTGAACATTAACATTAAGGTC 1864
Db 2656 ACAAAGTGAAGAAATGTTTATCAATGCTCTCAAGGGAACGTATACGTCAAGTGC 2715
Qy 1865 AAGCATATATGTGCGGTGGACCAAAAATCTCTGTTGGCAATTTGTGAATTA 1920
Db 2716 AGGCTTCAATGTAACGTAAGTCCGAAACCTTTCTTAAAGCATTTGATCAATTA 2771

RESULT 8
ADY33779
ID ADY33779 standard; DNA; 1305 BP.
XX
AC ADY33779;
XX
DT 05-MAY-2005 (first entry)
XX
DE Bacillus species alkaline protease coding sequence.
XX
KM mutagenesis; alkaline protease; detergent; laundry; cosmetic; food;
XX bleaching detergent; gene; ds.
OS Bacillus sp.
XX
FH Key
FT CDS location/qualifiers
FT 1..1305 /*tag= a
FT /product= "alkaline protease"
FT /note= "no start codon given"
FT /partial

PN EP1347044-A2.
XX
PD 24-SEP-2003.
XX
PF 21-MAR-2003; 2P03BP-00006472.
XX
PR 22-MAR-2002; 2002JP-00081428.
XX 06-JUN-2002; 2002JP-00165987.
PR 18-OCT-2002; 2002JP-00304230.
XX 18-OCT-2002; 2002JP-00304231.
XX
XX (KAOS) KAO CORP.
XX
PI Okuda M, Sato T, Saito K, Sumitomo N, Izawa Y, Saeki K;
PI Kobayashi T, Nomura M;
DR WPI; 2003-846540/79.
XX P-PSDB; ADY33778.
XX
PT New alkaline protease having specific amino acid residue at a specific
PT position of its amino acid sequence, useful for producing detergent
PT compositions, laundry detergent, fiber modifiers, leather-treating agents
PT or pipe cleaners.
XX
XX Disclosure; SEQ ID NO 2; 31pp; English.
XX
CC The invention relates to an alkaline protease having a fully defined
CC sequence of 434 amino acids (I) given in the specification, or an amino
CC acid sequence at least 80% homology with (I), where an amino acid residue
CC at position 65, 101, 163, 170, 171, 273, 320, 359 or 387 of (I) is
CC selected from 16 amino acid residues. The amino acid residues at the
CC corresponding positions are selected from: position 65: proline; position
CC 101: asparagine; position 163: histidine, aspartic acid, phenylalanine,
CC lysine, asparagine, serine, isoleucine, leucine, glutamine, threonine or
CC valine; position 170: valine or leucine; position 171: alanine, glutamic
CC acid, glycine or threonine; position 273: isoleucine, glycine or
CC threonine; position 320: phenylalanine, valine, threonine, leucine,
CC isoleucine or glycine; position 359: serine, leucine, valine, isoleucine
CC or glutamic acid; and position 387: alanine, lysine, glutamine, glutamic
CC acid, arginine or histidine. The alkaline protease is useful for the
CC production of a detergent composition, such as laundry detergent, fiber
CC modifiers, leather-treating agents, cosmetic compositions, bath
CC additives, food modifiers and pharmaceutical compositions. The alkaline
CC protease may also be used as bleaching detergent, hard surface cleaning
CC detergent, pipe cleaner, artificial tooth cleaner, and as a sterilizing
CC cleanser for medical tools. The new alkaline protease has a more potent
CC proteolytic capacity, exhibiting excellent detergency for the removal of
CC a complex stain, and has high secretion capacity. This sequence
CC corresponds to the *Bacillus* sp. alkaline protease coding sequence.
XX
XX Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;
SQ

Query Match 52.2%; Score 1002.6; DB 11; Length 1305;
Best Local Similarity 85.5%; Pred. No. 1.4e-305;
Matches 116; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

DB 241 TCACCTAATAAAGAAATGCGCCTCAGCGCAATCTAGTCTTCCATCTTATCATGATAGC 300
QY 916 AGTGTGGGCTTGAAGCTTGCCTTCCATCTGCAAACTTATTCAGCAAGCATTCAGT 975
DB 301 GGTGGGGGACTTGAAGACTTACCTTCCGATCTGCAAACTTATTCAGCAAGCATTCAGT 360
QY 976 GCAAGTCCCAAAATTCATCAAACTCTGCGGGGCGAGCGGTGAATGGGCTTACACAGCA 1035
DB 361 GCTGGGCGCAAAATTCATCAAACTCTGCGGGGCGAGCGGTGAATGGGCTTACACAA 420
QY 1036 GATTCAGAAATGTGGATGACTATGTAAGAAAAATATATGACATCTCTTTGGGGCT 1095
DB 421 GATTCAGAAATGTGGATGACTATGTAAGAAAAATATATGACATCTCTTTGGGGCT 480
QY 1096 GGAATGAAAGGCCGAACGGCGGTACCATCACTGCACTTGTAACGGCTAATAACGGCATA 1155
DB 481 GGAATGAAAGGCCGAACGGCGGTACCATCACTGCACTTGTAACGGCTAATAACGGCATA 540
QY 1156 ACAATGCGCGCAACCGAAACCTCGGTCACAGCTTCGGTTCCTATGCAATTAATTAAC 1215
DB 541 ACAATGCGCGCAACCGAAACCTCGGTCACAGCTTCGGTTCCTATGCAATTAATTAAC 600
QY 1216 CAGTTGCAAGTTCTTCCGTTGGCCCGCAAAAAGATGGGGAATCAAGCTGATGTC 1275
DB 601 CATGTGGCAAGTTCTTCCGTTGGCCCGCAAAAAGATGGGGAATCAAGCTGATGTC 660
QY 1276 ATGGCGCCAGGACATACATTTTATCAGCAAGATCTTCTTTCACCCGATTCCTCTTC 1335
DB 661 ATGGCAACGGGGAACGTTCACTATACAGAAATCTTCTTTCGACCGAATCTCTCTTC 720
QY 1336 TGGGGAATCATGACAGCAAAATATGCTATATGGGTGAAGCTCCATGCAACCGAAT 1395
DB 721 TGGGGAACCAAGACATGAATATATGCTATATGGGTGAAGCTCCATGCAACCGAAT 780
QY 1396 GTTGGGGGAATGTGACAGCTCCGTGAGCAATTTTGAATAAATAGAAATCACTCT 1455
DB 781 GTTGGGAACATGTGACAGCTCCGTGAGCAATTTTGAATAAATAGAAATCACTCT 840
QY 1456 AAGCCTTCCTTATGAAGCAAGCTTGTGATGCAAGGTGCTGATGTTGATGGGTAT 1515
DB 841 AAGCCTTCCTTATGAAGCAAGCTTGTGATGCAAGGTGCTGATGTTGATGGGTAT 900
QY 1516 CCGAACGGAACCAAGATGGGCGGAGTACCTTGATTAATCGTTGAACGTTGCTAT 1575
DB 901 CCGAACGGAACCAAGATGGGCGGAGTACCTTGATTAATCGTTGAACGTTGCTAT 960
QY 1576 GTGAAGCAATCAGAGCCCTATCACTAGCCCAAAAAGCAATTAACCTTTCCTGAAG 1635
DB 961 GTGAAGCAATCAGAGCCCTATCACTAGCCCAAAAAGCAATTAACCTTTCCTGAAG 1020
QY 1636 GCGGGCAAGCCATTGAATAATCTCCGTGATGTGGATGCCCTGCAAGCACTACTGCT 1695
DB 1021 GCGGGCAAGCCATTGAATAATCTCCGTGATGTGGATGCCCTGCAAGCACTACTGCT 1080
QY 1696 TCTGTAACTTGTGCAATGATTTGGATTTGTCATTACAGCAACAAAGCAAGATAT 1755
DB 1081 TCTGTAACTTGTGCAATGATTTGGATTTGTCATTACAGCAACAAAGCAAGATAT 1140
QY 1756 GTGGGAATGACTTCTGACACCAATTTGACATACTGGGATGGCGCAATAACGTAGAA 1815
DB 1141 GTGGGAATGACTTCTGACACCAATTTGACATACTGGGATGGCGCAATAACGTAGAA 1200
QY 1816 AATGATTTATTAATTCGCCCAAGTGAACATATACATTAAGAGTGCAACATATAAT 1875
DB 1201 AATGATTTATTAATTCGCCCAAGTGAACATATACATTAAGAGTGCAACATATAAT 1260
QY 1876 GTGGGTTGACCAAAAATCTTCTGTTGGCAATTTGAACATA 1920
DB 1261 GTACCGGTTGACCAAAAATCTTCTGTTGGCAATTTGAACATA 1305

RESULT 9

ADL25803
ID ADL25803 standard; DNA; 1305 BP.
XX
AC ADL25803;
XX
DT 20-MAY-2004 (first entry)
XX
DE Bacillus alkaline protease coding sequence.
XX
KW alkaline protease; washing agent; gene; ds; enzyme.
XX
OS Bacillus sp.
XX
FH Key location/Qualifiers
FT CDS 1..1305
FT /*tag=a
FT /partial
FT /product="Bacillus alkaline protease"
XX /note="No start codon is given"
XX
XX JP2004008085-A.
XX
XX 15-JAN-2004.
XX
XX 06-JUN-2002; 2002JP-00165950.
XX
XX 06-JUN-2002; 2002JP-00165950.
XX
XX (KAO) KAO CORP.
XX
XX MPI; 2004-094297/10.
XX
XX P-PSDB; ADL25802.
XX
XX Novel mutant alkaline protease produced by substituting the amino acid
XX residues useful as washing agent.
XX
XX Disclosure; SEQ ID NO 2; 21bp; Japanese.
XX
XX The invention comprises a mutant Bacillus alkaline protease which
XX contains substitutions at positions 163, 170 and 434. The mutant alkaline
XX protease is useful as a washing agent. The present DNA sequence encodes a
XX Bacillus alkaline protease of the invention.
XX
SQ Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;
Query Match 52.2%; Score 1002.6; DB 12; Length 1305;
Best Local Similarity 85.5%; Pred. No. 1.4e-305;
Matches 116; Conservative 0; Mismatches 189; Indels 0; Gaps 0;
QY 616 AATGATGCGCCAGATGTTGTCTCAACCGATGTGCGACAGACGCTACGGTTTGTAT 675
DB 1 AATATATGTTGGCGGTGAATGTCTCAACCGATGTGCGTCAAGACGCTACGGGTTGTAT 60
QY 676 GGACAAAGCCAGATTTGTGCGAGTTGCGATATCTGATTTGATACAGAAAGACGACAGT 735
DB 61 GGACAAAGACGATCTGACGCGTTGCCATACAGGCGTTGATACGATCGCATGACAGT 120
QY 736 TCGATGATGAAGCTTCCGCGGTAAATAACAGACATATATGCACTGGTTCGACGAT 795
DB 121 TCGATGATGAAGCTTCCGCGGGAATAATTAATGATATATATGCACTGGGACGACGAT 180
QY 796 AATGCGAATGATACGAACGCTTCATGATGATCCATGTCGACGAGTTCGATTTAGAAATGCG 855
DB 181 AATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 856 GCAAGAAATTAAGAAATGCAAGCTCAAGCATCTGTTTTTCATTCATCATCATGATAGC 915
DB 241 TCCACTAATTAAGAAATGCGCGCTTCAGCGCATCTAGCTTCCATCTCATCTGATAGC 300
QY 916 AGTGTGGCTTGAAGCTTCCATCTCAATCTGCAACCTTATTTAGCCAAAGCATTCAGT 975
DB 301 GGTGGGGGACTTGGAGGACTACCTTCGATCTGCAAACTTATTTAGCCAAAGCATTCAGT 360

QY 976 GCAGGTGCAGAAATTCATTAACAACTCCCTGCGGGGAGCGGTGAATGGGGCTTACACGACA 1035
DB 361 GCTGTGCGCAAAATTCATTAACAACTCTTGGGAGACGATGAATGGGGCTTACACACACA 420
QY 1036 GATTTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095
DB 421 GATTTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 1096 GGGAAATGAAGCCGGAACGGGGTACCATCTGATGATGATGATGATGATGATGATGATGAT 1155
DB 481 GGGAAATGAAGCCGGAACGGGGTACCATCTGATGATGATGATGATGATGATGATGATGAT 540
QY 1156 ACAGTGGCGGCAACGGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1215
DB 541 ACAGTGGCGGCAACGGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 1216 CACGTTCACAGTTCTCTTCCGCGGCGGACCAAAAGATGGGCAATCAACCTGATGATG 1275
DB 601 CATGTGGACAGTTCTCTTCCGCGGCGGACCAAAAGATGGGCAATCAACCTGATGATGATG 660
QY 1276 ATGGCGGCAAGGACATGATTTTATGACCAAGATCTTCTTTCGACCGGATTCCTCTTC 1335
DB 661 ATGGCGGCAAGGACATGATTTTATGACCAAGATCTTCTTTCGACCGGATTCCTCTTC 720
QY 1336 TGGGCGAATCATGACAGCAATATATGCTATATGAGTGGGAAGTGCATGAGCAACCGATT 1395
DB 721 TGGGCGAATCATGACAGCAATATATGCTATATGAGTGGGAAGTGCATGAGCAACCGATT 780
QY 1396 GTTGCGGGGAATGTTGACAGCTCCGTGAGCATTTTGTGAAAAATGAGGAATCATCTCT 1455
DB 781 GTTGCGGGGAATGTTGACAGCTCCGTGAGCATTTTGTGAAAAATGAGGAATCATCTCT 840
QY 1456 AAGCTTCTCTTATGAAAGCAGCTTTGATGATGATGATGATGATGATGATGATGATGAT 1515
DB 841 AAGCTTCTCTTATGAAAGCAGCTTTGATGATGATGATGATGATGATGATGATGATGAT 900
QY 1516 CCGAAGGGAACCAAGATGGGGGCGAGTGAACCTCGATTAATGATGATGATGATGATGAT 1575
DB 901 CCGAAGGGAACCAAGATGGGGGCGAGTGAACCTCGATTAATGATGATGATGATGATGAT 960
QY 1576 GTGAACGAATCAGTGCCTTATCACTAGCCAAAGGACATATTAATCTTATCTGCAACG 1635
DB 961 GTGAACGAATCAGTGCCTTATCACTAGCCAAAGGACATATTAATCTTATCTGCAACG 1020
QY 1636 GCGGGCAAGCCATTGAAATCTCCCTGATGATGATGATGATGATGATGATGATGATGAT 1695
DB 1021 GCGGGCAAGCCATTGAAATCTCCCTGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1696 TCTGTAACCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1755
DB 1081 TCTGTAACCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1756 GTCCGGAATGACTTCTGACACATTTGACAAATACTGGATGATGATGATGATGATGAT 1815
DB 1141 GTAGGAATGACTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1816 AATGATTTTATTAATGCGCCCAAGTGAACATATACATGATGATGATGATGATGAT 1875
DB 1201 AATGATTTTATTAATGCAACCAAGGCGGACGATATCAATGATGATGATGATGAT 1260
QY 1876 GTGCGGTTGGAACCAAACTCTCGTTGGCAATTTGAACTAA 1920
DB 1261 GTACCGGTTGGAACCAACCTCTCTGTTGCAATTTGAACTAA 1305
RESULT 10
ID AAQ27516 standard; DNA; 1299 BP.
XX
XX AAQ27516;
XX
XX DT 05-FEB-1993 (first entry)
XX

DB Alkali-protease Ya enzyme gene.
XX Alkali resistance, surface active agent resistance, detergency improver;
KW 88.
XX Bacillus sp. Y.
XX
FH Key Location/Qualifiers
PT CDS 1..1299
/tag= a
PN JP04197182-A.
XX
PD 16-JUL-1992.
XX
PF 28-NOV-1990; 90JP-00327110.
XX
PR 28-NOV-1990; 90JP-00327110.
XX
PA (LIOY) LION CORP.
XX
DR WPI; 1992-288440/35.
XX P-PSDB; AAR26274.
XX
PT DNA coding alkali-protease Ya enzyme - has good alkali and surfactant
XX resistance and improves detergency.
XX
PS Claim 3; Page 2; 17pp; Japanese.
XX
CC The sequence is that of the alkali-protease Ya enzyme gene which can be
CC used in the recombinant production of Ya enzyme. Ya enzyme is excellent
CC in alkali resistance and surface active agent resistance and improves
CC detergency
XX
SQ Sequence 1299 BP; 401 A; 246 C; 297 G; 355 T; 0 U; 0 Other;
Query Match 38.9%; Score 747.6; DB 2; Length 1299;
Beet Local Similarity 74.0%; Pred. No. 6.5e-225;
Matches 961; Conservative 0; Mismatches 334; Indels 3; Gaps 1;
QY 616 AATGATGTGGCCAGAGTATTGTCAGAGCGAGTGTGCGACAGAGAGTACGGTTGTAT 675
DB 1 AATGATGTGCAAGAGGAGTATGAAAGCTGATGTCACAAACCAATTACGATTATAT 60
QY 676 GGAAGAGCCAGATTTGTCAGTTGCCATGACTGATGATGATGATGATGATGATGAT 735
DB 61 GGAAGAGGTCAGTACTAGTTGTCAGTACGACAGGCTTATGATGATGATGATGAT 120
QY 736 TCGATTCATGAGCCCTTCCGCGGTAAATTAACAGCATATATGCACTGGTGGAGCAT 795
DB 121 TCTATGCAATGAGCAATTCGCGGGAATTAACAGGCTTATGATGATGATGATGAT 180
QY 796 AATGCAATGATATGCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 855
DB 181 AATGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 237
QY 856 GCAACGAATTAAGAAATGCACTCAAGCGAACTGTGTTTTCATCATCATCATCAT 915
DB 238 GCTTAAATTAAGAAATGCTCGCAAGCTTAACTTATGATGATGATGATGATGAT 297
QY 916 AGTGTGGGCTTGAAGGCTTGGCTTGCATTCGCAAACTTATTCAGCGAAGCATTCAG 975
DB 298 ACGGAGGATTTAGTGGCTTACCATCGAATTAAATACCTTATTTAGTCAAGCTTGAAT 357
QY 976 GCAAGTGCAGAAATTCATCAAACTCTGCGGGGCGAGCGGTGATGATGATGATGAT 1035
DB 358 GCTGAGAGCAAAATTCATCAAACTCTTGGGAGCCCGCAATTAATGAGGATGATGAT 417
QY 1036 GATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095
DB 418 AACTCGAGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 477
QY 1096 GGGAAATGAAGGCCCAAGCGCGGTACATCATGATGATGATGATGATGATGATGAT 1155

DB 478 GGTAAATGAAGTCTTAATTCAGAAACAAATTAATGCTCCAGATGACAGAAATGCTATT 537
QY 1156 ACAGTGGGCGCAACCAAAACCTGCGTCCAGGCTGCTTCCATGACATTAATTAAT 1215
DB 538 ACGGTGGGCGCAACGAAACCTATGCGCAAGCTTCCGTTGATGATGATGATGATGAT 597
QY 1216 CACGTTGCAAGTTCCTTCCGCGGCGGCAAGAAATGAGGGAATCAAGCTGATGTC 1275
DB 598 CATATTGCAAAATTTTCATGAGAGAGGCTACAGAGGATGACAAATTAAGCCTGACGTA 657
QY 1276 ATGGGCGCGGAGCATATATTTATGCAAGATCTTCTCTTGCACCGAATTCCTCTTC 1335
DB 658 ACAGCTCTGAAACATTTATTTATGACAGCTTCTCTTCTGCTCAGACTCTTGGTAT 717
QY 1336 TGGGCGAATCATGACAGCAAAATATGCTTATATGAGTGAACGTCATGCAACAGATT 1395
DB 718 TGGGCGAATTTATTAACGATTAATACGCGTATATGAGGCGTACCTCCATGCGACACTATT 777
QY 1396 GTTGGCGGGAATGTTGCAAGCTCCGTCAGCAATTTGTGAAAATAGAGAACTACTCCT 1455
DB 778 GTTGCAAGGAATGTCGCGCAATTAAGTGAAGCAATTTATTAATAATAGAGTATTACTCT 837
QY 1456 AAGCCTTCCCTATTTGAAAGCAGCTTGAATGACAGGTCGTCGATGATGATGATGAT 1515
DB 838 AAGCCTTCTTTATTAATAAGCTGACTTATGCTGCGTACGATGATGATGATGATGAT 897
QY 1516 CCGAAGCAAAACCAAGATGAGGCGGAGTGAACCTGATTAATGCTTGAACCTTGCCTAT 1575
DB 898 CTAAGTGGGACCAAGGCTGGGCGCTGTATCTGATTAATGCTTAAATGATGATGAT 957
QY 1576 GTGAACGAATTCAGTGCCTTATCAATGACCAAAAGCGACATATACCTTATCTGCAAG 1635
DB 958 GTCAATGAAGCAATGCAATTAAGCTCAAGCAAAAGCAACGATTCGTTCCAAAGCAAA 1017
QY 1636 GGGGCGAAGCCATTTGAAAATTCCTCGTATGATGATGATGATGATGATGATGATGAT 1695
DB 1018 GGGGCGAAGCCATTTGAAAATTCCTCGTATGATGATGATGATGATGATGATGATGAT 1077
QY 1696 TCTGTAAACCTGTCATATGATTTGATTTGATTCATTAACAGCAACAAACGAATAT 1755
DB 1078 TCTTATACCTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1137
QY 1756 GTGGGAATGACTTTCAGACCACTTTCATTAATTAATTAATTAATTAATTAATTAAT 1815
DB 1138 GTAGAAATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1197
QY 1816 AATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1875
DB 1198 AACGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1257
QY 1876 GTGCGGTTGACCAAAACCTTCTGTTGGCAATTTG 1913
DB 1258 GTACCATTCGCGCCACAGCGTTTTCATGACTGATGCT 1295
RESULT 11
ADM99124
ID ADM99124 standard; DNA, 1332 BP.
XX
AC ADM99124;
XX
DT 15-JUL-2004 (first entry)
XX
DB Bacterial source DNA encoding a protease Segid 7.
XX
KW gene; ds; protease; transgenic; detergent; fruit juice; syrup; puree;
KW paper; oral care; contact lens cleaning composition; hairball remedy;
XX antimicrobial.
OS Bacteria.
XX
FN WO2004033668-A2.

Db 709 KKAKESSYRBRBMYWKKMGTYRYYRWRSCMTBRAMSRRKRWAGASMSCWMYWRGARS 650
QY 814 GGTGATGTGATCCCATGTGGCAGTTGCGTATAGAAATGGCGCAAGCAATTAAGAAATG 873
Db 649 MWYSKYSGCAKCCCKRTMYTSSYSTMGYSYSKMSWTSMSTGMYTCTMTYTSMG 590
QY 874 GCACTCAGCAAGCAATCTGTTTTCATCATCATGATGATGAGTGGGCTTGAGAGC 933
Db 589 STRBSKMGWSMSRMYRWKMKMRKRKMYKMYKCTYRMCCTYRMTYTTTSSRSM 530
QY 934 TTGCTTCGAATCTGCAAACTTATTCAGCAAGCAATTCAGTGCAGTGCAGCAATTCAT 993
Db 529 YGARYARTSKRRYMYWYKRYKMYWYKMYKMSYMMRYGCRACKKCYAMCKKAY 470
QY 994 ACAAACCTCTGGGGGCGAGCGGTGAATGGGCTCAGAG 1032
Db 469 SGMMYMYRKYSKMMMSYTKYMSMYKCKRSMKYGANG 431

RESULT 13

ADA71938
ID ADA71938 standard; DNA, 2000 BP.

AC ADA71938;

DT 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 5263.

DE Plant; bacterial infection; fungal infection; viral infection; rice;

KW gene; ds.

XX Oryza sativa.

OS WO2003000898-A1.

XX 03-JAN-2003.

PD 22-JUN-2001; 2001WO-1B001105.

PR 22-JUN-2001; 2001WO-1B001105.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

DR WPI: 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899bp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 2.4%; Score 47; DB 8; Length 2000;

Best Local Similarity 11.2%; Pred. No. 0.0033;
Matches 102; Conservative 406; Mismatches 386; Indels 17; Gaps 5;

QY 1009 GCAAGCGGTGAATGGGCGCTACAGACAGATTCAGAAATGTGATGATGATGAAGAA 1068
Db 70 SBAASRTSSKSSSSGYSKGNKRYKRSKRWGRGRGRGRGRGRGRGRGRGRGRGR 129
QY 1069 AATGATATGACATCTTTTGGCGGCTGGAATGAAGCCGCAAGCGGTACATCACT 1128
Db 130 GSSGMMGCKSSMSYMMCYARGCSCKRKSSKGSWGTCTRGARGGSGMSGAKYKSG 189
QY 1129 GC--ACCTGTACGCTTAAAGCCATPAACAGTGGCGCAACCAAAACCTGCTCCA 1186
Db 190 SMSKMMMSSCGSCGGRSAVSRYGYSRKYKMYTYSASCMAYMTTYSWACS 249
QY 1187 GCTTGCTTCATGACATATATATATACAGAGTTCAGATTCCTCCGTCGCGCA 1246
Db 250 SYTTCRSKRBSMMMKMRKMSRYGYSRYKMYTYSASCMAYMTTYSWACS 309
QY 1247 CAAGAATGAGGCGATCAAGCCTGATGATG-----GCGCCAGGACATATATTTA 1299
Db 310 ATRYWGRYMSMAMMYKMYWYRGYKGMKGMWAGRAMMSMGRMSKACYMRMRM 369
QY 1300 TCAGCAATCTTCTTTCGACCCGATTCCTCTTCGGGCGAATCATGACGAAATAT 1359
Db 370 TFRRRWAKSSRSTRKRKRYKMYRMYGYSRMSCKRARMKCRSGRAMMGR 429
QY 1360 GCCTATATGGGTGAACGTCATGCGCAACCGATTCGCGGGAATGTCACAGCTC 1419
Db 430 GCMTCRMSYGMWRKMSWRMSKRYKMSRMYR--WRKKCSRTTMMGKTRGMMGTMR 487
QY 1420 CGTGAGCAATTTGTGAAATAGAGAAATCTCTCAAGCTTCCCTATTTGAAGCACT 1479
Db 488 CRKRSKMGKRCRRBRBMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 547
QY 1480 TTGATTCAGGTGCTGCTGATGTTGATGTTGTTATCCAGACGAAACCAAGATGGGC 1539
Db 548 GTYMAAMMKRYKRYMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYK 607
QY 1540 CGAGTACCCCTGGAATATGTTGAACGTTGCTATGTAAGCAATTCAGTCCCATCA 1599
Db 608 SAMSKMSRSCRKCRKASRSARRYAMMGTSGRMSRMSRYCTYMKMSMSTCTW 667
QY 1600 ACTAGCAAAAAGACATATATCTTTACTGCAAGCGGCGCAAGCAATTTGAAATCTCC 1659
Db 668 MYMSKTYAKKGSYMRIRYRAMCMYMRMYRMYRMYRMYRMYRMYRMYRMYRMYR 727
QY 1660 CTGTATGTCGATGATCCCTGCAAGCACTACTGCTTCTGTAACCTGCTCAATATTTG 1719
Db 728 RYMTSMYKCKSKMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 787
QY 1720 GATTTGCTTATACGAC-----CAACGGAACAAGATATGTCGGGAATGCTCTGAG 1774
Db 788 YMWYMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 847
QY 1775 CACCATTTGACATATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1833
Db 848 ARYMYKATATCATKMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 907
QY 1834 CCCCAAGTGAACATATATACATGATGATGATGATGATGATGATGATGATGATG 1893
Db 908 WKARWGSAAVYMMWKGAKMMWMMKGRMWGTYKMYVCTTMMKACBATTYMCAGMW 967
QY 1894 AACTTCGCTT 1904
Db 968 AMYSYSWTRY 978

RESULT 14

ACL37108
ID ACL37108 standard; cDNA, 2000 BP.

AC ACL37108;

DT 02-JUN-2005 (first entry)

XX

DE	Rice stress-regulated promoter SEQ ID NO:15671.
XX	
XX	88; abiotic stress tolerance; transgenic plant; plant; cereal;
KW	agriculture.
KW	
OS	Oryza sativa.
XX	
FN	WO2003008540-A2.
PD	
XX	30-JAN-2003.
XX	
PF	21-JUN-2002; 2002WO-US019668.
XX	
PR	22-JUN-2001; 2001US-0300112P.
XX	
PR	24-AUG-2001; 2001US-0314662P.
PR	26-SEP-2001; 2001US-0325277P.
PR	21-NOV-2001; 2001US-0332132P.
XX	
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.
XX	
P1	Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
PI	Moughamer T, Provart N, Rieke D, Zhu T;
XX	
DR	WPI; 2003-248011/24.
XX	
XX	New stress-responsive nucleic acid, useful for altering the
PT	responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT	stress, salt stress or osmotic stress.
XX	
PS	Claim 48; SEQ ID NO 15671; 89pp; English.
XX	
CC	The invention relates to novel abiotic stress responsive polynucleotides
CC	and polypeptides. Also disclosed are vectors, expression cassettes, host
CC	cells, and plants containing such polynucleotides. Also disclosed are
CC	methods for using the polynucleotides and polypeptides to alter the
CC	responsiveness of a plant to abiotic stress. The invention is useful in
CC	agriculture. The nucleic acid is useful for determining whether a test
CC	plant has been exposed to an abiotic stress condition. It is also useful
CC	for selecting an agent that alters abiotic stress regulated
CC	polynucleotide expression in a plant cell, and to identify a homolog or
CC	ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC	molecule and the polypeptide encoded by it are useful in altering the
CC	responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC	stress, osmotic stress or any of their combinations. The present sequence
CC	is used in the exemplification of the invention
XX	
SO	Sequence 2000 BP; 510 A; 350 C; 268 G; 518 T; 0 U; 354 Other;
	Query Match 2.1%; Score 40.6; DB 11; Length 2000;
	Best Local Similarity 11.6%; Pred. No. 0.35;
	Matches 46; Conservative 171; Mismatches 180; Indels 0; Gaps 0
OY	118 ATTCAAAACAACCGAGTGTGCTGGTTTCCTCAACACGCAACAAACAGTGCGCTGCA 177
DB :
	44 WYYTYCYKRWMTSMKYRAVMKRYMWMYKKMAAMAYACMMYAYGYMMKMYSHTWTY 103
OY	178 TTTCGTGTGAGTCTGAATAATGTGAACCTTCTTAAGATTGCTAAGAACACTTGAACA 237
DB :
	104 TRAYTKMKMRKRKMSRBAALAYTYSYMTSRKMYGYKKRMWRARRYYTWYCSTMS 163
OY	238 GTACCGGCAGAAATATAACTCATATGTGCCATTCAATGCGCCCATTTTAGAAGAAACA 297
DB :
	164 WNRWTRYWMMKYRAYKMAAYTTTTYTTCMMWWCTGYGKMKAAPASAKRRABAMMM 223
OY	298 AAACAGAACGTGAGACACCTGAGACAAAGATTCTGCATACATCCTCTGATTATGCATAT 357
DB :
	224 RWKSWTATAATATYSWMSMNAIWKRKRFKWGMWMAVSSTMMWRWWWYCTTRRAMYMKATWT 283
OY	358 ATTYGCGAGTATGAGGGGAGTGTCTAGTCAAAAGATCCGCTCATTTGAACACGAGATCA 417
DB :
	264 RCTWTMMMAAMTSANRAWKAGNATSATATKMYAKRTTYMKRRRYAYCAMATRTTKWSMM 343
OY	418 GTGAGGCCATCTTGCCGAATAATACAAATAGATCCCACGCTTTTCACAAAGCGCATCG 477

Db 344 RTKRWYTHARSQCKKMMWAAKRAATYCMATVSTYAYMRBRMRMSWMAAYSAKTY 403
 Qy 478 ACGCTGTGTAAGACGCTTGCGCTGTATACGAAGCA 514
 Db 404 YAKSYYYMYKCAKXWTTTAYYAAAABAKTRPAWGR 440
 RESULT 15
 ID AAS5654 standard; DNA; 522 BP.
 AC AAS5654;
 XX AAS5654;
 DT 19-DEC-2001 (first entry)
 DE DNA encoding novel subtilisin homologue #56.
 XX Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 KW contact lens cleansing solution; dry cleaning; ds.
 XX *Bacillus* sp.
 CS WO2001.75087-A2.
 XX
 PM 11-OCT-2001.
 PD
 XX
 PF 02-APR-2001; 2001WO-US010781.
 PR 03-APR-2000; 2000US-0194143P.
 XX
 PA (MAXY-) MAXYGEN INC.
 PI Nees JE, Welch M, Givler LD, Cherry JR, Borchert TV, Stemmer WPC;
 PI Minshull J;
 DR WPI; 2001-616689/71.
 XX P-PSDB; AAU38799.
 XX
 PT Novel subtilisin homologue polypeptides having improved endo-protease
 PT activity relative to mature Savinase subtilisin polypeptide, useful as
 PT component of cleaning solutions e.g. laundry detergents, dry cleaning.
 XX
 PS Claim 15; Page 82; 138pp; English.
 XX
 CC The invention relates to an isolated polypeptide with improved endo-
 CC protease activity relative to subtilisin homologue polypeptide, Savinase
 CC (RPM). The invention also relates to an integrated system comprising a
 CC computer or computer readable medium comprising a database comprising one
 CC or more sequence records. Each record comprises one or more character
 CC string corresponding to a nucleic acid or protein sequence of AAS56599-
 CC AAS56728 or AAU38744-AAU38873 respectively. The integrated system
 CC comprises a user input interface allowing a user to selectively enter or
 CC more sequence record. The integrated system is useful for presenting
 CC information pertaining to one of several sequence records stored in a
 CC database. The method involves determining a list of one or more character
 CC strings corresponding to the sequence or its subsequence, determining
 CC which character strings of the list are selected by a user, and
 CC displaying the selected character strings or aligning the selected
 CC character string with an additional character string. The method further
 CC involves displaying an alignment of the selected character string with
 CC the additional character string, and displaying the list. The
 CC polynucleotides are useful as probes, primers, sense and antisense
 CC nucleotides. The subtilisin homologue polypeptides are useful for
 CC producing antipodes which have diagnostic uses related to the activity,
 CC distribution and expression of subtilisin homologues. The subtilisin
 CC homologue polypeptides are used in compositions that serve as cleaning
 CC solutions in a wide variety of applications including laundry detergents,
 CC contact lens cleansing solutions, and dry cleaning. AAS56599-AAS56728
 CC represent the coding sequences of novel subtilisin homologues of the
 CC invention
 XX
 S0 Sequence 522 BP; 143 A; 124 C; 141G; 114 T; 0 U; 0 Other;

```

Query Match      2.1%; Score 40; DB 4; Length 522;
Best Local Similarity 72.2%; Pred. No. 0.25;
Matches 52; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy      1343 ATCATGACAGCAATATGCTTATATGGTGAACGTCCATGGCAACCCGATTGTGCGG 1402
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      447 ATCCGGGAGACATATGCTTCAATGAGCGGCAATCAATGGCAACCCGATTGTGCGG 506
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy      1403 GGAATGTTGCAC 1414
         | ||| |||
Db      507 GCGTGCAGCAC 518
         ||| |||
    
```

Search completed: April 8, 2006, 03:01:47
 Job time : 2983.67 secs

Y:

Y:

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 01:05:15 ; Search time 6935.08 Seconds
(without alignments)
12953.156 Million cell updates/sec

Title: US-10-784-870-3

Perfect score: 1930
Sequence: 1 atggaagaagaagaagctgctt.....cgttcgaactgtgaactaa 1920

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hnc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss81:*
10: gb_gss82:*
11: gb_gss83:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	47.8	2.5	997	10	CNS005TB
2	45	2.3	449	3	BJ359939
3	43.6	2.3	633	3	BJ369190
4	43.6	2.3	641	3	BJ393925
5	43.4	2.3	494	3	BJ334651
6	43.4	2.3	601	3	BJ387574
7	43	2.2	725	8	DR166768
8	41.6	2.2	634	5	BK253415
9	41	2.1	672	8	DN732737
10	41	2.1	1265	8	DN679903
11	41	2.1	1358	8	DN689766
12	40.8	2.1	678	10	CNS001981
13	40.4	2.1	854	8	DR951879
14	39.8	2.1	426	9	AZ145773
15	39.8	2.1	872	7	CO116672
16	39.8	2.1	1101	10	CNS00396
17	39.6	2.1	560	5	BJ55317
18	39.6	2.1	579	3	BM442028
19	39.6	2.1	696	3	BM442170
20	39.6	2.1	700	2	BF574449
21	39.6	2.1	710	5	BJ764471
22	39.6	2.1	721	10	CZ280243

C 23	39.6	2.1	737	5	BO764315
C 24	39.6	2.1	929	6	CA973725
C 25	39.6	2.1	1211	4	BC029393
C 26	39.6	2.1	1577	4	BC012590
C 27	39.4	2.1	1030	8	DN703987
C 28	39.4	2.1	1234	8	DN709335
C 29	39.2	2.0	466	3	BM442304
C 30	39.2	2.0	559	10	CM049955
C 31	39.2	2.0	599	10	CM112692
C 32	39.2	2.0	600	9	BZ360416
C 33	39.2	2.0	608	9	BZ364026
C 34	39	2.0	936	10	CL109289
C 35	38.8	2.0	695	1	A1729280
C 36	38.8	2.0	897	10	CL102311
C 37	38.8	2.0	1101	10	CNS00370
C 38	38.6	2.0	423	9	AQ437507
C 39	38.6	2.0	451	9	AZ745756
C 40	38.6	2.0	655	9	BH082837
C 41	38.6	2.0	717	3	BJ335636
C 42	38.4	2.0	629	10	CM118700
C 43	38.4	2.0	634	10	CE723448
C 44	38.4	2.0	690	3	B1698380
C 45	38.4	2.0	732	8	DN201554

ALIGNMENTS

RESULT 1
CNS005TB 997 bp DNA linear GSS 03-JUN-1999
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence T813 end of BAC # BACR12K22 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL060767
VERSION
AL060767.1 GI:4943573

KEYWORDS
Drosophila melanogaster (fruit fly)

SOURCE
Drosophila melanogaster

ORGANISM
Bukariyota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 997)

AUTHORS
TITLE
JOURNAL

COMMENT

Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Ooegawa and Aaron Mammox in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. 997
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR12K22"
/clone_id="RPCI-98"
/note="end : T813"

ORIGIN

National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source
1. .641
/organism="Dictyostellium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dd83105"
/sex="mat A"
/dev_stage="Slug stage"
/clone_1ib="Dictyostellium discoideum cDNA library, SF"

ORIGIN

Query Match 2.3%; Score 43.6; DB 3; Length 641;
Best Local Similarity 61.4%; Pred. No. 0.54; Mismatches 44; Indels 0; Gaps 0;
Matches 70; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 1205 ATATATATTAACCAAGTTCACAGTCTCTTCCCGTGGCCGACAAAGATGGCGGAATCA 1264
|||||
DB 169 ATATGAAATATATATGAGTATCTTCATCAAAAGGTCCACACATGATGTAGATTTGA 228

QY 1265 AGCGTATGTCAAGGCGCCGAGACATACATTTTATCAGCAAGATCTTCTCTTG 1318
|||||
DB 229 AACCTGATATATGTTGACCTGTGGATATATATTCATCGCGCAAGATCAATGCTG 282

RESULT 5

LOCUS BJ334651 494 bp mRNA linear EST 05-MAR-2002
DEFINITION BJ334651 Dictyostellium discoideum cDNA library, AF Dictyostellium
discoideum cDNA clone dda47f08 5', mRNA sequence.
ACCESSION BJ334651
VERSION BJ334651.1 GI:19164781
KEYWORDS EST.

SOURCE

Dictyostellium discoideum
Dictyostellium discoideum
Bukaryota; Mycetozoa; Dictyostellida; Dictyostellium.

REFERENCE 1 (bases 1 to 494)
Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-1, T.
Full length cDNA of Dictyostellium discoideum at the aggregation
stage
Unpublished (2002)

JOURNAL

Unpublished (2002)
Contact: Tadaeu Shin-1
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source
1. .494
/organism="Dictyostellium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dda47f08"
/sex="mat A"
/dev_stage="Aggregation stage"
/clone_1ib="Dictyostellium discoideum cDNA library, AF"

ORIGIN

Query Match 2.3%; Score 43.4; DB 3; Length 494;
Best Local Similarity 60.7%; Pred. No. 0.57; Mismatches 46; Indels 0; Gaps 0;
Matches 71; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1203 AGATATATTAACCAAGTTCACAGTCTCTTCCCGTGGCCGACAAAGATGGCGGAAT 1262
|||||
DB 244 ATACATATGAAATATATATTTGTTTCATTCATCAAAAGGTCCACACATGATGTAGAT 303

QY 1263 CAAGCTGATGTCATGGCCGAGGACATACATTTTATACGAAGATCTTCTCTG 1319
|||||
DB 304 GAACCTGATATATGTTGCCCTGTGGATATATATTCATCGCGCAAGATCAATGCTG 360

RESULT 6

LOCUS BJ387574 601 bp mRNA linear EST 08-MAR-2002
DEFINITION BJ387574 Dictyostellium discoideum cDNA library, SF Dictyostellium
discoideum cDNA clone dda3a18 5', mRNA sequence.
ACCESSION BJ387574
VERSION BJ387574.1 GI:19296958
KEYWORDS EST.

SOURCE

Dictyostellium discoideum
Dictyostellium discoideum
Bukaryota; Mycetozoa; Dictyostellida; Dictyostellium.

REFERENCE 1 (bases 1 to 601)
Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-1, T.
Full length cDNA of Dictyostellium discoideum at the slug stage
Unpublished (2002)

JOURNAL

Contact: Tadaeu Shin-1
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source
1. .601
/organism="Dictyostellium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dda3a18"
/sex="mat A"
/dev_stage="Slug stage"
/clone_1ib="Dictyostellium discoideum cDNA library, SF"

ORIGIN

Query Match 2.3%; Score 43.4; DB 3; Length 601;
Best Local Similarity 60.7%; Pred. No. 0.61; Mismatches 46; Indels 0; Gaps 0;
Matches 71; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1203 AGATATATTAACCAAGTTCACAGTCTCTTCCCGTGGCCGACAAAGATGGCGGAAT 1262
|||||
DB 47 ATACATATGAAATATATATTTGTTTCATTCATCAAAAGGTCCACACATGATGTAGAT 106

QY 1263 CAAGCTGATGTCATGGCCGAGGACATACATTTTATCAGCAAGATCTTCTCTG 1319
|||||
DB 107 GAACCTGATATATGTTGCCCTGTGGATATATATTCATCGCGCAAGATCAATGCTG 163

RESULT 7

LOCUS DR166768/c 725 bp mRNA linear EST 20-JUN-2005
DEFINITION DR166768 RTPHOS1_14_G08_b1_A029 roots minus phosphorous Pinus taeda cDNA
clone RTPHOS1_14_G08_A029 3', mRNA sequence.
ACCESSION DR166768
VERSION DR166768.1 GI:68026324
KEYWORDS EST.

SOURCE

Pinus taeda (loblolly pine)
Pinus taeda
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE

1 (bases 1 to 725)
Pratt, L., Cordonnier-Pratt, M.M., Lorenz, W.W., Zimmermann, C. and
Dean, J.F.D.
An EST database from phosphorous-deficient loblolly pine (Pinus
taeda) roots
Unpublished (2005)

JOURNAL

Other ESTs: RPHOS1_14_G08_GL_A029
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 533 0310

Email: mmpatt@uga.edu
RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia); plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FBRC) and the CCONES project at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement as they have been trimmed to exclude polyA.
Seq primer: M13-21 (TGTAAACACGCGCCAGT)
POLYA=No.

FEATURES

Source

Location/Qualifiers

1. .725
/organism="Pinus taeda"
/mol_type="mRNA"
/strain="3 CLONES"
/db_xref="taxon:3152"
/clone="RTPH01_14_G08_A029"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_1lb="Roots minus phosphorus"
/note="Organ: Root; Vector: pSL180; Site: 1: EcoRI;
Site_2: XhoI; The library was prepared from polyA+ RNA
from the roots of 1-year-old loblolly pine (*Pinus taeda*)
cuttings that were rooted and then planted in washed sand.
The rooted cuttings were maintained for 19 days (July 28
2003 harvest) under ambient conditions in a local
greenhouse. They were kept on a weekly regimen of 0.5x
nutrient-complete Hoagland's solution and supplemented
with additional water sufficient to maintain a 15% soil
moisture content. Starting five weeks prior to harvesting
roots for mRNA preparation, the weekly fertilizer regimen
for the potted trees was modified to omit phosphorus from
the 0.5x Hoagland's solution (5 applications). For two
days prior to starting the modified fertilizer regimen,
plants were flushed extensively with water to reduce
residual levels of phosphorus. Double-stranded cDNA was
cloned unidirectionally into pSL180. Inserts can be
excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match	2.2%	Score 43;	DB 8;	Length 725;
Best Local Similarity	49.8%;	Pred. No. 0.85;		
Matches 109;	Conservative 0;	Mismatches 110;	Indels 0;	Gaps 0;

Qy	CGCGCTGATATGCAAGGAACTTTTGGATCTGGATTTTAAAGAAATTCACAAACAACCCATG	138
Db	CGGCTATCTCTGTATTAAGCTGCTGATCGGCACGCTCAATGCAAAAAGCACCAACCTGG	451
Qy	TCAGTGGTTTCTCCAAACAGCGCAAAACAGTGGCGCTGCATTTCTGTGGAGTCTGAAA	186
Db	AAACAATCAAGAGACAACTAAAGGAGCTTCGGAGCTTGGCAATGGCAATGGCTCGAA	391
Qy	ATGTGAACTTTCTTAAGGATGCTTAAGAACTTGAACACGTAACGGCAATATATTAAC	256
Db	AATTTAAGCTTCTTCACTGAGCTGGAAGACATATAAAGCAGACCTTGCAATTTGCCAAG	331
Qy	TCCATATTGTCCAATTCATGGCCCATTTTAAGAGAAA	295
Db	AGCGTTGTACACATTAAGAGAGAAAATAAAGTACTAA	292

RESULT 8	LOCUS	DEFINITION	ACCESSION
BX253415	BX253415	634 bp mRNA linear EST 24-FEB-2007	BX253415
		BX253415 Pinus pinaster differentiating xylem adult pinaster	
		cDNA clone PF083C05, mRNA sequence.	
			BX253415

VERSION	BX253415.1	GI:28520153
KEYWORDS	EST.	
SOURCE	pinus pinaster	
ORGANISM	Pinus pinaster	

REFERENCE
1 (bases 1 to 634)
Canton, F.R., Le Provost, G., Garcia, V., Barre, A., Frigerio, J.-M.,

TITLE	Transcriptome analysis of wood formation in maritime pine
JOURNAL	Unpublished (2003)
COMMENT	Contact: Prigierio JM

FEATURES

Source

Location/Qualifiers

```

/organism="Pinus pinaster"
/mol_type="rRNA"
/strain="ecotype: Corsican"
/db_xref="taxon:71647"
/clone="pP083C05"
/tissue_type="differentiating xylem"
/dev_stage="adult"
/clone.lib="Uni-Zap XR lambda (Stratagene) ; Site 1: Eco
RI, Site 2: Xho I; A composite cDNA library was made with
mRNA isolated from normal, compression, opposite, early
and late wood of Maritime pine uni-directionally cloned
into Uni-Zap XR using the SP(-) cDNA Synthesis kit
(Stratagene). pBluescript SK(-) plasmids were obtained by
in vivo mass excision. The nucleotide sequence with the T3
primer by GENOME EXPRESS, Meylan, France"

```

ORIGIN

Query Match	2.2%	Score 41.6	DB 5	Length 634
Best Local Similarity	49.1%	Pred. No. 2.1		
Matches 110; Conservative	0	Mismatches 114	Indels 0	Gaps 0;

Qy	137	TCAGTGGTTTCTCCAAACAGCGACAAACAGGTGGCGGTGCAATTTCTGTGAGTCTGGAAA	196
Db	81	CAGCTATCTCTGTAAAGTGTCTGATGTGGCACTGCAATGCAAAAAGCAACCAACGTGG	140
Qy	77	CGGCTGGTATCCAAAGAACTTTATCTTCGAAATTTTAAAGGAATTCAAACAACACGATG	136
Db	141	AACCAATCAAGAGACAACTTAAAGGATCTCCGGAAGTTGGGATGGCAATGGCTGGCAA	200
Qy	197	ATGTGAACTTTTAAAGGATTGCTTAAAGAACTTGAAACGATGCGGCAATATATTAAC	256
Db	201	AAGTTAAGCTTCTTACGTGAGCTGGAAGACATTTAAAGCAACCTTGCAATTTGCCAAG	260
Qy	257	TCCATATGTGCAATTCATGGCCCCATTTTAGAAGAAACAAA	300
Db	261	AGGTTGTACCAATTAGAGAGAAAATTAAGTAAAGGAA	304

RESULT 9	
DN733737	
LOCUS	DN732737
DEFINITION	672 bp mRNA linear EST 31-MAR-2005
	CNB66-A05.xid-t SHGC-CNB Gasferosteus aculeatus cDNA clone
	CNB66-A05 3', mRNA sequence.

SOURCE ORGANISM	Source
<i>Gasterosteus aculeatus</i> (three spined stickleback)	1
<i>Gasterosteus aculeatus</i>	2
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;	3

REFERENCE
1 (bases 1 to 672)
Kingsley, D.M., Petchel, C., Balabhadra, S., Grimwood, J., Dickson, M.,
Schmutz, J., and Myers, R.M.
Expressed sequence tags from *Gasterosteus aculeatus*
Unpublished (2003)
JOURNAL
COMMENT
Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@shgc.stanford.edu
Plate: 66
High quality sequence start: 23
High quality sequence stop: 672.
Location/Qualifiers

FEATURES

Source

1..672
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Conner Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CMB6-405"
/sex="mixed male and female"
/tissue_type="brain"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_11b="SHGC-CNB"
/note="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligo(dT) sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTGTTCTAGATCGCGAGCGCGCC(T)25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5 prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxxATC' (where is ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at:
http://www.openbiosystems.com/cdna_library_construction_fa_q.php#8 The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria. Clones available from Open Biosystems:
http://www.openbiosystems.com/stickleback"

ORIGIN

Query Match 2.1%, Score 41; DB 8; Length 672;
Best Local Similarity 49.8%; Pred. No. 3.3;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 85 GATGCAAGACTTTGATCTGGATTAAAGAAATTCACAAACGAGTCAAGTGT 144
DB 258 GAGACACAGAGTCTTCATTTTCACATCAGTCTTTTCAAAAAGAGGCCCCCT 317
QY 145 TTCTCAAAACGACGCAACAGTGCCTGTCATTTCTGTGAGCTGAAATGTGAAA 204
DB 318 GCCCCCCCCCCCAAGCAAGAACTATAGTCAACAAGTTCACACAAAGTATTA 377
QY 205 CTCTTAAAGATGCTAAAGAACTTGAACAGTACCGCAATATTAATTCATTT 264
DB 378 TTACATCAACAATCCCAAGATATAGAAATGATACAAAGCTTAAAGTCACTTG 437
QY 265 GTCCCAATCAATGAGCCCATTTTGAAGA 293
DB 438 GTGCAATCGTGCCCTTTTAAAGAGA 466

RESULT 10
DN679903

LOCUS
DN679903 1265 bp mRNA linear EST 30-MAR-2005
DEFINITION
CGX08-G08 x1d-t SHGC-CGX *Gasterosteus aculeatus* cDNA clone
CGX08-G08 3', mRNA sequence.
ACCESSION
DN679903
VERSION
DN679903.1 GI:62022153
KEYWORDS
EST.
SOURCE
ORGANISM
Gasterosteus aculeatus (three spined stickleback)
Gasterosteus aculeatus
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; *Gasterosteiformes*;
Gasterosteidae; *Gasterosteus*.
1 (bases 1 to 1265)
REFERENCE
Kingsley, D.M., Petchel, C., Balabhadra, S., Grimwood, J., Dickson, M.,
Schmutz, J., and Myers, R.M.
Expressed sequence tags from *Gasterosteus aculeatus*
Unpublished (2003)
JOURNAL
COMMENT
Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@shgc.stanford.edu
Plate: 08
High quality sequence start: 15
High quality sequence stop: 593.
Location/Qualifiers

FEATURES

Source

1..1265
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Conner Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CGX08-G08"
/sex="mixed male and female"
/tissue_type="eyes"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_11b="SHGC-CGX"
/note="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligo(dT) sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTGTTCTAGATCGCGAGCGCGCC(T)25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5 prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxxATC' (where is ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at:
http://www.openbiosystems.com/cdna_library_construction_fa_q.php#8 The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria. Clones available from Open Biosystems:
http://www.openbiosystems.com/stickleback"

ORIGIN

Query Match 2.1%, Score 41; DB 8; Length 1265;
Best Local Similarity 49.8%; Pred. No. 4;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 85 GATGCAAGACTTTGATCTGGATTAAAGAAATTCACAAACGAGTCAAGTGT 144
DB 204 GAGACACAGAGTCTTCATTTTCACATCAGTCTTTTCAAAAAGAGGCCCCCT 263
QY 145 TTCTCAAAACGACGCAACAGTGCCTGTCATTTCTGTGAGCTGAAATGTGAAA 204
DB 264 GCCCCCCCCCCCAAGCAAGAACTATAGTCAACAAGTTCACACAAAGTATTA 323

Oy	205	CTTCTTAAGAGTTGCTAAAGCAACTTGAACAGTACCGCCAAATATATAACCTCCATAT	264
Oy	324	TTACATCAACAAATCCCAAGTTAAATGAATATATACAGCTTTAAAGTCTTACGCTTG	383
Oy	265	GTCCAAATTCATAGGCCCATTTTAAAGA	293
Db	384	GTGATATCGGTGCGCTTCTTTTAAAGA	412

RESULT 11	DN689766	1358 bp	mRNA	linear	EST 30-MAR-2005
LOCUS	DN689766				
DEFINITION	CGX66-E11.xid-t SHGC-CGX Gasterosteus aculeatus cDNA clone				
ACCESSION	DN689766				
VERSION	DN689766.1	GI:62042478			
KEYWORDS	EST.				
SOURCE	Gasterosteus aculeatus (three spined stickleback)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes; Gasterosteidae; Gasterosteus.				
REFERENCE	1 (bases 1 to 1358)				
AUTHORS	Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M., Schmutz,J. and Myers,R.M.				
TITLE	Expressed sequence tags from Gasterosteus aculeatus				
COMMENT	Unpublished (2003)				
	Contact: Grimwood, Jane				
	Stanford Human Genome Center				
	Stanford University School of Medicine				
	975 S California Ave, Palo Alto, CA 94304, USA				
	Tel: 650 320 5917				
	Fax: 650 320 5801				
	Email: jane@hgc.stanford.edu				
	Plate: 66				
	High quality sequence start: 39				
	High quality sequence stop: 687.				
FEATURES	Location/Qualifiers				
source	1..1358				
	/organism="Gasterosteus aculeatus"				
	/mol_type="mRNA"				
	/strain="Conner Creek sticklebacks, WA USA"				
	/db_xref="taxon:69293"				
	/clone="CGX66-E11"				
	/sex="mixed male and female"				
	/tissue_type="eyes"				
	/dev_stage="adult"				
	/lab_host="DH10B (T1 phage resistant)"				
	/clone_11b="SHGC-CGX"				
	/note="Vector: Express 1, Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligodT sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTAGTTCTAATCGCAGCGCGCCCTT25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5 prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxATTC' (where 'x' is ATC is the second half of the EcoRV site, and 'xxx' is derived from the cDNA sequence). A map of the Express 1 vector is available at: http://www.openbiosystems.com/cdna_library_constructionfaq.php#8 The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria. Clones available from Open Biosystems: http://www.openbiosystems.com/stickleback"				

Query Match

2.1%; Score 41; DB 8; Length 1358;

Query Match	2.1%	Score 40.8	DB 10	Length 678
ORIGIN				
Query Match	2.1%	Score 40.8	DB 10	Length 678

Best Local Similarity 53.8%; Pred. No. 3.8;
Matches 84; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 187 GAGTCTGAAATGTAAGTCTTAAGATGCTTAAGAACTTAACAGTACCGCA 246
 DB 498 GAGAAAGAAAAGAAAAGTCAATTAAGAAATGAGACAACTTCGAAACCTCCGAGC 439
 QY 247 AATAATTAATCTCATTTGCTCAATTCATGCGCCCATTTTGAAGAAACAAACGAAG 306
 DB 438 AATATTAAGCGCAGCACTTCAATCAATGAGAGTCCAGAGAGAGAGAGAGAGAGC 379
 QY 307 CTGAGACAACTGAGCAAGATTCGACTACATC 342
 DB 378 CATGAGAAAATACTTGAGAGATTAATAGTTGAAAAC 343

RESULT 13
 DR951879 854 bp mRNA linear EST 02-AUG-2005
 LOCUS ESR1143418 Aquilegia cDNA library Aquilegia formosa x Aquilegia
 DEFINITION pubescens cDNA clone COIRX48, mRNA sequence.
 ACCESSION DR951879
 VERSION DR951879.1 GI:71721242
 KEYWORDS EST.
 SOURCE Aquilegia formosa x Aquilegia pubescens
 ORGANISM Aquilegia formosa x Aquilegia pubescens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
 Ranunculaceae; Aquilegia.
 1 (bases 1 to 854)
 Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E.,
 Nordborg, M., and Tomkins, J.
 Generation of ESTs from Aquilegia
 Unpublished (2005)
 Other ESTs: ESR1143417
 Contact: Scott Hodges
 Department of Ecology, Evolution and Marine Biology
 University of California, Santa Barbara
 Santa Barbara, CA 93106, USA
 Tel: 805 893 7813
 Fax: 805 893 4724
 Email: shodges@lifesci.ucsb.edu
 Seq primer: M13 Reverse.
 FEATURES
 SOURCE Location/Qualifiers
 1..854
 /organism="Aquilegia formosa x Aquilegia pubescens"
 /mol_type="mRNA"
 /db_xref="taxon:338618"
 /clone="COIRX48"
 /tissue_type="mixed shoot and floral apical meristems,
 flower buds, leaves and roots"
 /lab_host="DH10B T1 (T1 and T5 phage resistance)"
 /clone_1lb="Aquilegia cDNA library"
 /note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;
 P2, P3, and P4 lines of Aquilegia formosa x A. pubescens
 were grown from seed in greenhouses at UC Santa Barbara.
 From these plants three sets of tissue were collected: 1)
 Small flower buds (<10 mm) and very young inflorescences
 (71 & 29% by weight respectively), 2) Medium (7-20 mm) and
 large (at or near anthesis) flower buds (65 & 35% by
 weight respectively) and 3) Shoot apical meristems. A
 fourth set of tissue was collected from plants of A.
 formosa. These plants were grown from seed in sand and at
 approximately 1 month root tissue and leaf tissue of
 various developmental stages were collected (84 & 16% by
 weight respectively). Total RNA was extracted from each
 set of tissue and pooled in the following proportions:
 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled
 total RNA, mRNA was extracted and enriched for full-length
 messages and then normalized with proprietary methods by
 Invitrogen."

ORIGIN

Query Match 2.1%; Score 40.4; DB 8; Length 854;
 Best Local Similarity 58.2%; Pred. No. 5.4;
 Matches 71; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 285 TTGAGAGAAACAAACAGAGCTAGAGCACTGAGCAAGATTTCTGACTACATCCC 344
 DB 403 TTCTGAGAGAGAAAAGATTAATCAAGAGAGATGAAGTAAAGTTAGGATTCGACGCA 462
 QY 345 TGATATGATATATTTGCTGAGTATGAGGGGAGTTCAGTCAAAAGTCCGTCATGCA 404
 DB 463 TGATGATGATGATGATGATGAGAGAGATGATGATTTGATTTGCAATTCACAGAGA 522
 QY 405 AC 406
 DB 523 TC 524

RESULT 14
 AZ145773 426 bp DNA linear GSS 28-AUG-2000
 LOCUS SP_0019_B1_H10_SP68 Strongylocentrotus purpuratus, purple sea
 urchin, sperm genomic BAC library Strongylocentrotus purpuratus
 genomic clone Plate=19 Col=19 Row=P, genomic survey sequence.
 ACCESSION AZ145773
 VERSION AZ145773.1 GI:8297676
 KEYWORDS GSS.
 SOURCE Strongylocentrotus purpuratus
 ORGANISM Strongylocentrotus purpuratus
 Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 Echinoidae; Euechinoidae; Echinacea; Echinoidae;
 Strongylocentrotidae; Strongylocentrotus.
 1 (bases 1 to 426)
 Cameron, R.A., Mahatras, G., Raat, J.P., Martinez, P., Biondi, T.R.,
 Swartzell, S., Wallace, J.C., Pousetka, A.J., Livingston, B.T.,
 Wray, G.A., Eitensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H.,
 and Hood, L.
 A sea urchin genome project: Sequence scan, virtual map, and
 additional resources
 Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)
 10920195
 Contact: Cameron, RA, Davidson, EH, Hood, L
 Division of Biology 156-29
 California Institute of Technology
 Pasadena California 91125, USA
 Tel: (626) 395-8421
 Fax: (626) 793-3047
 Email: acameron@caltech.edu
 Plate: 19 row: P column: 19
 Seq primer: SP6
 Class: BAC ends
 High quality sequence step: 426.
 FEATURES
 SOURCE Location/Qualifiers
 1..426
 /organism="Strongylocentrotus purpuratus"
 /mol_type="genomic DNA"
 /db_xref="taxon:7668"
 /clone="Plate=19 Col=19 Row=P"
 /clone_1lb="Strongylocentrotus purpuratus, purple sea
 urchin, sperm genomic BAC library"
 /note="Organ: sperm; Vector: BACs; 6; BAC clones in E-Coli
 DH10B"

ORIGIN

Query Match 2.1%; Score 39.8; DB 9; Length 426;
 Best Local Similarity 62.6%; Pred. No. 6.5;
 Matches 62; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 439 TACAAAATAGATCCCAAGCTTTTCAAAAAGCGCATGACGCTGTGAAGCGCTGGCG 498
 DB 102 TACCTTTTGGTGTCTATATATTCACAAAAGTGACATTCACATGCTTAAACATTTGCA 161
 QY 499 CTGATACGAGACGAAACATTAAGAGTGCATTAAGA 537

Db 162 CTCGATTCATTCAGTGCATATAAGAGCTCAGCAGACA 200

RESULT 15

COL16672

LOCUS

DEFINITION GR_EB019A19.r GR_Eb Gossypium raimondii cDNA clone GR_EB019A19

ACCESSION

COL16672

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

Search completed: April 8, 2006, 08:09:32
Job time : 6938.08 secs

Query Match 2.1%; Score 39.8; DB 7; Length 872;
Best Local Similarity 59.1%; Pred. No. 8.2; 47; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

1391 CGATTGTCGCGGAGATGTCACAGCTCCGTGAGCATTTGTGAAAAATGAGGATCA 1450
|||||
416 CTATTGGTGTGGAACATATTCCTGAAATGTTGCTTCTTAATAAAAAATATGAAAAA 475
|||||

1451 CTCCTAAGCCCTTCCCTATTGAAGAGCCTGATTGACAGTGCCTGATGTTGG 1505
|||||
476 GTCCTGACAGCTTCCTATTGAAAGAGGTTCTTTCAGTGCATGCTGCGCATGG 530
|||||

ORIGIN

1. .872
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_EB019A19"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_id="GR_Eb"
/note="Vector: pCMVSPORT-6.1; Site 1: NotI; Site 2: EcoRV; library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGT. More glycerol clones held in -80."

FEATURES

source

1. .872
location/Qualifiers

1. .872
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_EB019A19"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_id="GR_Eb"
/note="Vector: pCMVSPORT-6.1; Site 1: NotI; Site 2: EcoRV; library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGT. More glycerol clones held in -80."

Kim H., Yu Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and Wing, R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 019 row: A column: 19.

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW nucleic - nucleic search, using sw model

Run on: April 8, 2006, 00:12:00 ; Search time 589.04 Seconds
(without alignments)
5794.037 Million cell updates/sec

Title: US-10-784-870-3

Perfect score: 1920
Sequence: 1 atgagaagaagaagggtgtt.....cgttgcaatgtgaactaa 1920

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/1_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6_COMB.seq:*
5: /cgn2_6/prodata/1/ina/6_COMB.seq:*
6: /cgn2_6/prodata/1/ina/6_COMB.seq:*
7: /cgn2_6/prodata/1/ina/6_COMB.seq:*
8: /cgn2_6/prodata/1/ina/6_COMB.seq:*
9: /cgn2_6/prodata/1/ina/6_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1920	100.0	1920	US-09-509-814A-3	Sequence 3, Appl
2	1920	100.0	1920	US-09-920-954-3	Sequence 3, Appl
3	1477.2	76.9	1923	US-09-509-814A-5	Sequence 5, Appl
4	1477.2	76.9	1923	US-09-509-814A-7	Sequence 7, Appl
5	1477.2	76.9	1923	US-09-920-954-5	Sequence 5, Appl
6	1477.2	76.9	1923	US-09-920-954-7	Sequence 7, Appl
7	1136.8	59.2	3003	US-08-873-479-41	Sequence 41, Appl
8	58.2	3.0	7218	US-08-332-463-14	Sequence 14, Appl
9	41.2	2.1	2001	US-09-949-016-2801	Sequence 2801, Ap
c 10	41	2.1	1141	US-09-806-708B-22	Sequence 22, Appl
11	40	2.1	522	US-09-824-893A-56	Sequence 56, Appl
12	40	2.1	522	US-09-824-893A-70	Sequence 70, Appl
13	40	2.1	1494	US-09-134-000C-2585	Sequence 2585, Ap
14	39.6	2.1	2947	US-09-949-016-815	Sequence 815, App
15	39	2.0	522	US-09-824-893A-50	Sequence 50, Appl
16	38.4	2.0	522	US-09-824-893A-37	Sequence 37, Appl
17	38.4	2.0	522	US-09-824-893A-43	Sequence 43, Appl
18	38.4	2.0	522	US-09-824-893A-47	Sequence 47, Appl
19	38.4	2.0	522	US-09-824-893A-60	Sequence 60, Appl
20	38.4	2.0	522	US-09-824-893A-62	Sequence 62, Appl
21	38.4	2.0	522	US-09-824-893A-66	Sequence 66, Appl
22	38.4	2.0	522	US-09-824-893A-76	Sequence 76, Appl
23	38.4	2.0	522	US-09-824-893A-79	Sequence 79, Appl
24	38.4	2.0	522	US-09-824-893A-80	Sequence 80, Appl

25	38.4	2.0	522	US-09-824-893A-83	Sequence 83, Appl
26	38.4	2.0	522	US-09-824-893A-101	Sequence 101, App
27	38.4	2.0	522	US-09-824-893A-112	Sequence 112, App
28	38.4	2.0	522	US-09-824-893A-118	Sequence 118, App
29	38.4	2.0	522	US-09-824-893A-119	Sequence 119, App
30	38.4	2.0	522	US-09-824-893A-120	Sequence 120, App
31	38.4	2.0	522	US-09-824-893A-122	Sequence 122, App
32	38.4	2.0	522	US-09-824-893A-125	Sequence 125, App
33	38.4	2.0	522	US-09-824-893A-130	Sequence 130, App
34	38.4	2.0	522	US-09-824-893A-130	Sequence 130, App
35	38.2	2.0	76810	US-09-949-016-13528	Sequence 12528, A
36	37.2	1.9	399	US-09-621-976-8976	Sequence 8976, Ap
37	36.8	1.9	522	US-09-824-893A-38	Sequence 38, Appl
38	36.8	1.9	522	US-09-824-893A-40	Sequence 40, Appl
39	36.8	1.9	522	US-09-824-893A-46	Sequence 46, Appl
40	36.8	1.9	522	US-09-824-893A-46	Sequence 46, Appl
41	36.8	1.9	522	US-09-824-893A-51	Sequence 51, Appl
42	36.8	1.9	522	US-09-824-893A-52	Sequence 52, Appl
43	36.8	1.9	522	US-09-824-893A-53	Sequence 53, Appl
44	36.8	1.9	522	US-09-824-893A-55	Sequence 55, Appl
45	36.8	1.9	522	US-09-824-893A-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-09-509-814A-3
; Sequence 3, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKATA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKI, KATSUHIKA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1920)
US-09-509-814A-3
Query Match 100.0%; Score 1920; DB 3; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAAAGAGAGAGTGTATCTGTATATCAGCTGACGATTCGTGACTGT 60
Db 1 ATGAGAAAGAGAGAGTGTATCTGTATATCAGCTGACGATTCGTGACTGT 60
QY 61 GATTAAACATCCCTCGCTGCTGATGACAGAGATTTGATCTGATTTAAAGATT 120
Db 61 GATTAAACATCCCTCGCTGCTGATGACAGAGATTTGATCTGATTTAAAGATT 120
QY 121 CAAGCAAGCAGATGTCAGTGTCTTCCAAAGCAGCAAAAGAGTGGCGCTCATTT 180
Db 121 CAAGCAAGCAGATGTCAGTGTCTTCCAAAGCAGCAAAAGAGTGGCGCTCATTT 180

QY 181 CTGGTGAAGTCTGAAAATGTGAACTTCTTAAAGATTGCTTAAAGAACTTGAACAGTA 240
DB 181 CTGGTGAAGTCTGAAAATGTGAACTTCTTAAAGATTGCTTAAAGAACTTGAACAGTA 240
QY 241 CCGGCAAAATTAATACTCCATATTGTCCAATTCATAGGCCCCCATTTTAAAGAAACAAA 300
DB 241 CCGGCAAAATTAATACTCCATATTGTCCAATTCATAGGCCCCCATTTTAAAGAAACAAA 300
QY 301 CAGAAAGCTAGAGACAACTGAGCAAAAGATTCTGACTACATCCCTGATTATGCAATATAT 360
DB 301 CAGAAAGCTAGAGACAACTGAGCAAAAGATTCTGACTACATCCCTGATTATGCAATATAT 360
QY 361 GTGCGATTAGAGGGGGATGTTGACTCAAAAGTCGCTCCATTGAAACAGCTGGAATCAGTG 420
DB 361 GTGCGATTAGAGGGGGATGTTGACTCAAAAGTCGCTCCATTGAAACAGCTGGAATCAGTG 420
QY 421 GAGCCATACCTTCCGAAATACAAATAGATCCCAAGCTTTTTCACAAAGGCGCATCGACG 480
DB 421 GAGCCATACCTTCCGAAATACAAATAGATCCCAAGCTTTTTCACAAAGGCGCATCGACG 480
QY 481 CTGGTGAAGCGTTGGCGCTTGTATACGAAAGCAACATTAAGAAAGTGCAATTAAGAGGC 540
DB 481 CTGGTGAAGCGTTGGCGCTTGTATACGAAAGCAACATTAAGAAAGTGCAATTAAGAGGC 540
QY 541 ATCGAAGAAATGCTCAGTACGTAAGCAAGCAATGACGTCCATTATATTACGGCAAGCCT 600
DB 541 ATCGAAGAAATGCTCAGTACGTAAGCAAGCAATGACGTCCATTATATTACGGCAAGCCT 600
QY 601 GAATATAAGGTATGATGATATGTCGACAGAGTATTTTCMAAGCGGATGTGCAACAGAC 660
DB 601 GAATATAAGGTATGATGATATGTCGACAGAGTATTTTCMAAGCGGATGTGCAACAGAC 660
QY 661 AGCTACGTTTGTATAGACAAAGCCAGATTGTGCGAGTTGCCGATATCTGATATGATACA 720
DB 661 AGCTACGTTTGTATAGACAAAGCCAGATTGTGCGAGTTGCCGATATCTGATATGATACA 720
QY 721 GGAAGAAACGACAGTTGATGATGATGAAAGCTTCCGCGGTAAATTAACAGCATATATGCA 780
DB 721 GGAAGAAACGACAGTTGATGATGATGAAAGCTTCCGCGGTAAATTAACAGCATATATGCA 780
QY 781 CTGGGTGCGACGAATATGCGAATGATACGAACGCTCATGATGATGATGATGATGATGAT 840
DB 781 CTGGGTGCGACGAATATGCGAATGATACGAACGCTCATGATGATGATGATGATGATGAT 840
QY 841 GTATTAGGAAATGCGGCAACGAATTAAGAAATGCGAATCTGCAAGCGGATCTGATTTTCA 900
DB 841 GTATTAGGAAATGCGGCAACGAATTAAGAAATGCGAATCTGCAAGCGGATCTGATTTTCA 900
QY 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 AGCCAAAGCATTCAGTGCAGAGTGCAGAAATTCATACAACTCCTGGGGGGGCGATGAT 1020
DB 961 AGCCAAAGCATTCAGTGCAGAGTGCAGAAATTCATACAACTCCTGGGGGGGCGATGAT 1020
QY 1021 GGGGCTTACAGACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 GGGGCTTACAGACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ATTTCTTTTCGCGGCTGGAATGAAGAGCGGAGACCGGATACCATGATGATGATGATGATGAT 1140
DB 1081 ATTTCTTTTCGCGGCTGGAATGAAGAGCGGAGACCGGATACCATGATGATGATGATGATGAT 1140
QY 1141 GCTAATAAACCCCATTAACAGTGCAGACCGAAGAACTGCGTCAAGCTTGGTTCAT 1200
DB 1141 GCTAATAAACCCCATTAACAGTGCAGACCGAAGAACTGCGTCAAGCTTGGTTCAT 1200
QY 1201 GCAAGTATATTAACAGTGTGCAAGTTCCTTTCCTGCGGCGGCAAAAGATGGGCGA 1260
DB 1201 GCAAGTATATTAACAGTGTGCAAGTTCCTTTCCTGCGGCGGCAAAAGATGGGCGA 1260

QY 1261 ATCAAGCTGATGTCATGCGCGACAGGACATACATTTTATCAGCAAGATCTTCTTCA 1320
DB 1261 ATCAAGCTGATGTCATGCGCGCGACAGGACATACATTTTATCAGCAAGATCTTCTTCA 1320
QY 1321 CCGGATTCCTCTTCTGCGGCAATCATGACAGCAATATGCTATATAGGTGGAACGTCC 1380
DB 1321 CCGGATTCCTCTTCTGCGGCAATCATGACAGCAATATGCTATATAGGTGGAACGTCC 1380
QY 1381 ATGCGACACCGAATTTGTTGCGGGGAATGTTGACAGCTCCGTTAGCATTTTGTAAAAAT 1440
DB 1381 ATGCGACACCGAATTTGTTGCGGGGAATGTTGACAGCTCCGTTAGCATTTTGTAAAAAT 1440
QY 1441 AGAGGAATCACTCTTCAAGCTTCCCTATTGAAACAGCTTGTGATGAGTGCCTGAT 1500
DB 1441 AGAGGAATCACTCTTCAAGCTTCCCTATTGAAACAGCTTGTGATGAGTGCCTGAT 1500
QY 1501 GTTGGATTGGGTTATCCGAAACGAAACCAAGATGAGGCGAGTGAACCTGGAATTAATCG 1560
DB 1501 GTTGGATTGGGTTATCCGAAACGAAACCAAGATGAGGCGAGTGAACCTGGAATTAATCG 1560
QY 1561 TTGAACGTTGCTTATGTAAGCAATCCAGTGCCTATCACTAGCCAAAAGCGACATAT 1620
DB 1561 TTGAACGTTGCTTATGTAAGCAATCCAGTGCCTATCACTAGCCAAAAGCGACATAT 1620
QY 1621 ACCTTACTGCAACGCGCGGCAAGCCATGAAATCTCCCTGATGAGTGCCTGAT 1680
DB 1621 ACCTTACTGCAACGCGCGGCAAGCCATGAAATCTCCCTGATGAGTGCCTGAT 1680
QY 1681 GCAAGCACTACTGCTTCTGTAACCTGCTCAATGATTTGATTTGTCATTAACGACCA 1740
DB 1681 GCAAGCACTACTGCTTCTGTAACCTGCTCAATGATTTGATTTGTCATTAACGACCA 1740
QY 1741 AACGGAACAATATATGTCGGAATATGATCTTCACACACATTTGCAATTAATGGAAGGC 1800
DB 1741 AACGGAACAATATATGTCGGAATATGATCTTCACACACATTTGCAATTAATGGAAGGC 1800
QY 1801 CGCATATAAGTAAATATGATTTATTAATTCGCCCCCAAGTGAACATATACCATTTAG 1860
DB 1801 CGCATATAAGTAAATATGATTTATTAATTCGCCCCCAAGTGAACATATACCATTTAG 1860
QY 1861 GTGCAAGCATTAATGTCGCGGTTGACACCAAAATCTTCTGTTGCAATTTGTAACATA 1920
DB 1861 GTGCAAGCATTAATGTCGCGGTTGACACCAAAATCTTCTGTTGCAATTTGTAACATA 1920

RESULT 2
US-09-920-954-3
; Sequence 3, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1920
; TYPE: DNA

ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
US-09-920-954-3

Query Match 100.0%; Score 1920; DB 3; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ATGAGAAAGAAAGAGTGTATTTATCTGTTTATCAAGCTGACAGATTCCTGCTGACTGTT 60
DB 1 ATGAGAAAGAAAGAGTGTATTTATCTGTTTATCAAGCTGACAGATTCCTGCTGACTGTT 60
OY 61 GCATTAAACATTCCTCGCTGCTGATGCAAGACCTTTGATCTGGATTTTAAAGAAATT 120
DB 61 GCATTAAACATTCCTCGCTGCTGATGCAAGACCTTTGATCTGGATTTTAAAGAAATT 120
OY 121 CAACCAACACGATGCTGAGTGTCTCCAAACAGCGACAAACAGGTGCGCTGATTT 180
DB 121 CAACCAACACGATGCTGAGTGTCTCCAAACAGCGACAAACAGGTGCGCTGATTT 180
OY 181 CTGGTGAAGTCTGAAAATGTGAACCTTCTTAAAGATTGCTAAAGAACTTGAACAGTA 240
DB 181 CTGGTGAAGTCTGAAAATGTGAACCTTCTTAAAGATTGCTAAAGAACTTGAACAGTA 240
OY 241 CCGGCAAAATTAATTAATCTCCATATTTGCTCAATTCATATGCGCCCATTTTGAAGAACAA 300
DB 241 CCGGCAAAATTAATTAATCTCCATATTTGCTCAATTCATATGCGCCCATTTTGAAGAACAA 300
OY 301 CAGAGCTAGAGCAACCTGAGCAAGATTTCTGACTACATCCCTGATTAATGCAATATTT 360
DB 301 CAGAGCTAGAGCAACCTGAGCAAGATTTCTGACTACATCCCTGATTAATGCAATATTT 360
OY 361 GTGAGTATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCATATGAAACGTGAAATCAGTG 420
DB 361 GTGAGTATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCATATGAAACGTGAAATCAGTG 420
OY 421 GAGCCATATCTTCCGAAATTAACAAATTAATCCCACTTTTCACAAAAGCGCATGAGC 480
DB 421 GAGCCATATCTTCCGAAATTAACAAATTAATCCCACTTTTCACAAAAGCGCATGAGC 480
OY 481 CTGGTGAAGAGCGTGGCGCTTGTATGAAAGCAAGAACTTAAGAAAGTGAATTAAGAGC 540
DB 481 CTGGTGAAGAGCGTGGCGCTTGTATGAAAGCAAGAACTTAAGAAAGTGAATTAAGAGC 540
OY 541 ATCGAGGAATCGCTCAAGTACGTAAGCAACAAATGACGTCTCAATTAATTAATGCGCAAGCCT 600
DB 541 ATCGAGGAATCGCTCAAGTACGTAAGCAACAAATGACGTCTCAATTAATTAATGCGCAAGCCT 600
OY 601 GAAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 GAAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
OY 661 AGCTACGCTTGTATGAGCAAGGCAAGATTTGTCAGATTTGTCAGATTTGTCAGATTTGTCAG 720
DB 661 AGCTACGCTTGTATGAGCAAGGCAAGATTTGTCAGATTTGTCAGATTTGTCAGATTTGTCAG 720
OY 721 GGAAGAAACGACGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 GGAAGAAACGACGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
OY 781 CTGGGTCCGACGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CTGGGTCCGACGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
OY 841 GATATTGGAATGGGCAACGATTAAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 900
DB 841 GATATTGGAATGGGCAACGATTAAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 900
OY 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
```

```
OY 961 AGCCAAAGCATTCAGTCCAGGTGCGAGAAATTCATCAAACTCTCGGGGGGCAAGCGTGAAT 1020
DB 961 AGCCAAAGCATTCAGTCCAGGTGCGAGAAATTCATCAAACTCTCGGGGGGCAAGCGTGAAT 1020
OY 1021 GGGGCTTACACGACAGATTCAGAAATGAGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 GGGGCTTACACGACAGATTCAGAAATGAGATGATGATGATGATGATGATGATGATGATGAT 1080
OY 1081 ATTTCTTTGCGCGCTGGGAATGAAAGCGCGGATACATCATGTCACCTGCTGATG 1140
DB 1081 ATTTCTTTGCGCGCTGGGAATGAAAGCGCGGATACATCATGTCACCTGCTGATG 1140
OY 1141 GCTAAAGACCCATTAACATGCTCGCGCAACCGAAACCTGCTGCAAGCTTGGTCTAT 1200
DB 1141 GCTAAAGACCCATTAACATGCTCGCGCAACCGAAACCTGCTGCAAGCTTGGTCTAT 1200
OY 1201 GCGATTAATTAATCAAGTTCAGATGCTCTCTCCGTTGCGCGCAAAAGATGCGGCA 1260
DB 1201 GCGATTAATTAATCAAGTTCAGATGCTCTCTCCGTTGCGCGCAAAAGATGCGGCA 1260
OY 1261 ATCAAGCTGATGATCAATGCGCGCAGGGAATTAATTAATGAGCAAGATCTTCTTGA 1320
DB 1261 ATCAAGCTGATGATCAATGCGCGCAGGGAATTAATTAATGAGCAAGATCTTCTTGA 1320
OY 1321 CCGATTCCTCTCTGCGGCAATCAATGACAGCAATATGCTATATGAGTGAACGTCC 1380
DB 1321 CCGATTCCTCTCTGCGGCAATCAATGACAGCAATATGCTATATGAGTGAACGTCC 1380
OY 1381 ATGGCAACCGATTTGCGGGGAAATGTTGACAGCTCCGAGAGATTTTGAAGAAAT 1440
DB 1381 ATGGCAACCGATTTGCGGGGAAATGTTGACAGCTCCGAGAGATTTTGAAGAAAT 1440
OY 1441 AGAGGAATCACTCTTAAGCTTCCCTATTAAGAAAGCTTTGATTTGAGAGTCTGAT 1500
DB 1441 AGAGGAATCACTCTTAAGCTTCCCTATTAAGAAAGCTTTGATTTGAGAGTCTGAT 1500
OY 1501 GTTGAATGGGTTATCCGAACGAAACCAAGGATGGGCGGAGTCACTGATTAATCG 1560
DB 1501 GTTGAATGGGTTATCCGAACGAAACCAAGGATGGGCGGAGTCACTGATTAATCG 1560
OY 1561 TTGAACGTTGCTTAATGTAACGAATCCAGTCCCTATCAATGACCAAAACGACATAT 1620
DB 1561 TTGAACGTTGCTTAATGTAACGAATCCAGTCCCTATCAATGACCAAAACGACATAT 1620
OY 1621 ACCTTTACTGCAACGGCGGCAAGCAATTAAGAAATCTCCCTGATGATGATGATGATGAT 1680
DB 1621 ACCTTTACTGCAACGGCGGCAAGCAATTAAGAAATCTCCCTGATGATGATGATGATGAT 1680
OY 1681 GCAAGCACTACTGCTTGTGAACCTGTGCAATGATTTGATTTGATTTGATTTGATTTGAT 1740
DB 1681 GCAAGCACTACTGCTTGTGAACCTGTGCAATGATTTGATTTGATTTGATTTGATTTGAT 1740
OY 1741 AACGGAACGAATATGTTGGGAATGATCTTCAAGCACTTTGAATTAATCTGGGATGCG 1800
DB 1741 AACGGAACGAATATGTTGGGAATGATCTTCAAGCACTTTGAATTAATCTGGGATGCG 1800
OY 1801 CGCAATTAACGTAAGAAATGATTTATTAATTTGATTTGATTTGATTTGATTTGATTTGAT 1860
DB 1801 CGCAATTAACGTAAGAAATGATTTATTAATTTGATTTGATTTGATTTGATTTGATTTGAT 1860
OY 1861 GTGCAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
DB 1861 GTGCAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
```

RESULT 3
US-09-509-814A-5
; Sequence 5, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI

Db 1750 AACGATATGTGGAATGACTTACTTCCGCAATACATGATTAACGTGGATGCCCAT 1809
Qy 1807 AACGTGAATAATATTATTATTAATTCGCCCAAGTGAACATATACATTGAGGTGCA 1866
Db 1810 AACGTGAATAATATTATTATTAATTCGCCCAAGTGAACATATGAGGTGACAG 1869
Qy 1867 GCATTAATGTGCGGCTTGACCAAACTTCTCGTTGGCAATTGTGAACCTAA 1920
Db 1870 GCTTAATGATGATCGGCTTGACCAAACTTCTCGTTGGCAATTGTGAATTA 1923

RESULT 4
US-09-509-814A-7
Sequence 7, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OR INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent in version 3.0
SEQ ID NO 7
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (1923)
US-09-509-814A-7

Query Match 76.9%; Score 1477.2; DB 3; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Qy 7 AAGAGAGGAGGTTTATTCGTTTATTCAGCTGACGCACTTCTGCGACCTTGTCATTA 66
Db 10 AAGAGAGGAGGTTTATTCGTTTATTCAGCTGACGCACTTCTGCGACCTTGTCATTA 69
Qy 67 AACATCCCTCGGCTGATGATCAAGCACTTTGATTCGATTTTAAAGAAATTCAAACA 126
Db 70 AGTATTCATCGCAGGTGATGATCAAGCACTTTGATTCGATTTTAAAGAAATTCAGACA 129
Qy 127 ACAACGATGTCAGTGTGTTCTTCACAAAGCAAAACAGTGTGCGGTGATTTCTGTG 186
Db 130 ACAACGATGTCAGTGTGTTCTTCACAAAGCAAAACAGTGTGCGGTGATTTCTGTG 189
Qy 187 GAGTGTGAATAATGTAACTTCTTAAGGATTCGTAAGAACTTGAACATACCGGCA 246
Db 190 GATGTGAATAATGTAACTTCTTAAGGATTCGTAAGAACTTGAACATACCGGCA 249
Qy 247 AATAATTAATCTCATATTGTCATTCATTCATGCGCCCATTTTGAAGAAACAAACAGAG 306
Db 250 AATAATTAATCTCATATTGTCATTCATTCATGCGCCCATTTTGAAGAAACAAACAGAG 309
Qy 307 CTAGAGCAACTGAGCAAGAAATTCCTGACATTCATTCCTGATTAATGATTAATGTCAG 366
Db 310 CTGAGCAAAACAGGGGCAAAAGATTCCTGACATTCATTCCTGATTAATGATTAATGTCAG 369
Qy 367 TATGAGGGGAGATTCAGTCAAAAGTCCGCTCATTTGAACAGTGAATCAGTGAAGCA 426

Db 370 TATGAGGGGAGATTCAGTCAAAAGTCCGCTCATTTGAACAGTGAATCAGTGAAGCTT 429
Qy 427 TACTTCCCAATATACAAATAGATCCCACTTTTTCACAAAGCCGATTCAGCGTGTG 486
Db 430 TATTTCCGATATACGAATATAGATCCCACTTTTTCACAAAGCCGATTCAGCGTGTG 489
Qy 487 AAAGCTTGGCGCTTATACGAAGCAAAATTAAGAAATGTCATTAAGAGGATTCAG 546
Db 490 AAAGCTTGGCGCTTATACGAAGCAAAATTAAGAAATGTCATTAAGAGGATTCAG 549
Qy 547 GAAATCGCTCAGTACGATGCAAGCAATGACGTCATTAATTAACGCAAAAGCTGAAT 606
Db 550 CAATTCGCAAAATTCGCAATTAAGCAATGATGCTATATTAACGCAAAAGCTGAT 609
Qy 607 AAGGTGATGAATGATGTCGCAAGGATTTGTCAAAGCCGATGTCAGAGCACTAC 666
Db 610 AAGGTGATGAATGATGTCGCGGTGGAATTTGTCAAAGCCGATGTCAGAGCACTAC 669
Qy 667 GGTGTGATGCAAAAGCCGATTTGTCAGATTCGCGATTAATGATTAACAGGAAGA 726
Db 670 GGTGTGATGCAAAAGCCGATTTGTCAGATTCGCGATTAATGATTAACAGGTCGC 729
Qy 727 AACGACATTCATGCAATGCAAGCCCTTCCGCGTAAATTAACGCACTAATGCACTGG 786
Db 730 AATGACATTCATGCAATGCAAGCCCTTCCGCGTAAATTAACGCACTAATGCACTGG 789
Qy 787 CGGAGCAATTAATGCAATGATACGAACGTCATGTCACCATGTCAGGTCGATTA 846
Db 790 CGGAGCAATTAATGCAATGATACGAATGTCATGTCACCATGTCAGGTCGATTA 849
Qy 847 GGAATATGCGGCAAGATTAAGAAATGCACTTCAAGCCGAATCTGATTTTCAATCATC 906
Db 850 GGAATATGCGGTCATTAATTAAGAAATGCACTTCAAGCCGAATCTGATTTTCAATCATC 909
Qy 907 ATGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
Db 910 ATGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 969
Qy 967 GCATTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1026
Db 970 GCATTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1029
Qy 1027 TACACGACGATTCGAGAAATGTGATGATGATGATGATGATGATGATGATGATGAT 1086
Db 1030 TACACGACGATTCGAGAAATGTGATGATGATGATGATGATGATGATGATGATGAT 1089
Qy 1087 TTGCGGCTGCGGAAATGAAGAGCCGAAACGCGGTACATCAATGTCAGTGCAGTGC 1146
Db 1090 TTGCGGCTGCGGAAATGAAGAGCCGAAACGCGGTACATCAATGTCAGTGCAGTGC 1149
Qy 1147 AACGCAATTAAGTGCAGGCGCAACCGAAACCTGTCGCAAGCTTGGTTCATGATGAT 1206
Db 1150 AATGCAATTAAGTGCAGGCGCAACCGAAACCTGTCGCAAGCTTGGTTCATGATGAT 1209
Qy 1207 AATATTAACAGTTCGACAGTTCCTTCCGTCGCGCCGCAAAAGATGAGGAAATCAAG 1266
Db 1210 AATATTAACAGTTCGACAGTTCCTTCCGTCGCGCCGCAAAAGATGAGGAAATCAAA 1269
Qy 1267 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
Db 1270 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1329
Qy 1327 TCTTCTCTGCGGCAATTCATGACGAATATGCTTATATGAGTGAAGAGTCCATGCA 1386
Db 1330 TCTTCTCTGCGGCAATTCATGACGAATATGCTTATATGAGTGAAGAGTCCATGCA 1389
Qy 1387 AACCGAATTTGTCGCGGAAATGTCAGAGCTCCGTCGAGCAATTTGTGAATAATGAGGA 1446
Db 1390 AACCGAATTTGTCGCGGAAATGTCAGAGCTCCGTCGAGCAATTTGTGAATAATGAGGA 1449
Qy 1447 ATCACTCTAAGCCCTTCCCTATTAAGAGCACTTGAATGAGGTCGTCGATGATGAG 1506

Db 1450 ATCAACCAAGGCTTCTCTATTAAAGGGCACTGATTGCCGGTCACTGACATCGGC 1509
Qy 1507 TTGGGTATCCGAACGAAACCAAGGATGGGGCCGAGTGACCTGATTAATCGTTGAAC 1566
Db 1510 CTGGCTACCCGAAACGATTAACAAAGATGGGGACAGATGAATGATTAATCCCTGAAC 1569
Qy 1567 GTTGCCTATGTAAGAAAGATCCAGTGGCCCTATCACTAGCAAAAAGGCAATATACCTTT 1626
Db 1570 GTTGCCTATGTAAGAAAGATCCAGTGGCCCTATCACTAGCAAAAAGGCAATATACCTTT 1629
Qy 1627 ACTGCAACGGCGGGCAAGCCATTTGAAAATCTCCGTATGCTGGATGCCCTGCAAGC 1686
Db 1630 ACTGCTATCGCCGGCAAGCTTTGAAAATCTCCGTATGCTGGATGCCCTGCGAGC 1689
Qy 1687 ACTACTGCTTCTGTAACTCTGCTCAATGATTTGGATTGGTCATTACAGCAACAAACGGA 1746
Db 1690 ACAACTGCTTCCGTAAACGTTGCTCAATGATCTGGAACCTTGTCAATACCGCTCAATAGGC 1749
Qy 1747 ACAAGATATGTCGGGAATGACTTCAAGCAACATTTGAATACCTGGGATGGCGCAAT 1806
Db 1750 ACAAGATATGTAAGAAATGAATTAATCTTCCATACATGATTAATACCTGGGATGGCGCAAT 1809
Qy 1807 AACGTAGAAATGATTTTATTAATTCGCCCAAGTGAACATATACATGAGGTGCA 1866
Db 1810 AACGTAGAAATGATTTTATTAATTCACCAACAAAGCGGAGGTATACATGAAATGATACAG 1869
Qy 1867 GCATTAATGTCGCCGTTGACCAAAAATCTTCTGTTGGCAATTGTGACTTA 1920
Db 1870 GCATTAATGTCGCCGTTGACCAAAAATCTTCTGTTGGCAATTGTGACTTA 1923

RESULT 5

US-09-920-954-5
Sequence 5, Application US/09920954
Patent No. 6759228

GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUYOSHI
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
PRIOR APPLICATION NUMBER: 2001-08-03
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP-9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-09-920-954-5

Query Match 76.9%; Score 1477.2; DB 3; Length 1923;
Best Local Similarity 85.7%; Pred.No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Qy 7 AAGAAAGAGGTGTTTATCTGTTTATCAAGTCAAGGATTTCTGTGACTGTTGACTTA 66
Db 10 AAGAAAGAGGTGTTTATCTGTTTATCAAGTCAAGGATTTCTGTGACTGTTGACTTA 69

Qy 67 AACATCCCTGGCTGGTATGCAAGACTTTTGTATCTGATTTTAAAGAAATTCAAACA 126
Db 70 AGTATCATCTGCAAGGTGGTCAAGAAATTTTATCTGATTTTCAAGAAATTCAGACA 129
Qy 127 ACAACCAATGTCAGTGGTTTCTCCAAACAGCAAAAACAGGTGGGCTGACTTTCTGGTG 186
Db 130 ACAACTATGCTAAAGTTTCTCCAAAGAGGGGAGACTGATGCTGCTGCTTTCTGGTG 189
Qy 187 GAGTCTGAATAATGTGAACCTTCTTAAGAAATTTGCTTAAGAAATTTGAACATTCGCGCA 246
Db 190 GAATCTGAATAATGTGAACCTTCTTAAGAAATTTGCTTAAGAAATTTGAACATTCGCGCA 249
Qy 247 AATATAAATCCCAATTTGCTCAATTCATTAAGTGGCCCAATTTTGAAGAAACAAACAGAG 306
Db 250 AATATAAATCCCAATTTGCTCAATTCATTAAGTGGCCCAATTTTGAAGAAACAAACAGAG 309
Qy 307 CTAGACAACTGAGCAAAAGATTTCTGACTACATCCCTGATTAATGATATATTTGTCAG 366
Db 310 CTGAAAGAAACAGGGGCAAAAGATTTCTGACTACATCTGATTAATGCTTACATTTGTGAG 369
Qy 367 TATGAGGGGATGTTCACTCAAGTCAAAAGTCCCTCAATGAAACGTTGAATCGTGAAGCA 426
Db 370 TATGAGGGGATGTTCACTCAAGTCAAAAGTCCCTCAATGAAACGTTGAATCGTGAAGCT 429
Qy 427 TACTTGGCAAAATCAAAATGATCCCAAGCTTTCAAAAGGCGCATGAGCTGGTG 486
Db 430 TATTGGCGAATACAGAAATGATCCCAAGCTTTCAAAAGGCGCATGAGCTGGTG 489
Qy 487 AAACGTTGGCGCTTGTATACCAACAGAAACATTAAGAAAGTGAATTAAGAGGATCGAG 546
Db 490 AAACGTTGGCGCTTGTATACCAACAGAAATTAAGAAAGTGAATTAAGAGGATCGAA 549
Qy 547 GAATTCCTCACTACATGACCAAGCAATGAGTCAATTAATTAAGCAAAAGCTGATAT 606
Db 550 CAATTCCTCACTACATGACCAAGCAATGAGTCAATTAATTAAGCAAAAGCTGATAT 609
Qy 607 AAGGTATGAATGATGTCGACAGAGTATGTCAAAGCGAGTGGCAAGAGAGCTAC 666
Db 610 AAGGTATGAATGATGTCGACAGAGTATGTCAAAGCGAGTGGCTCAAGAGAGCTAC 669
Qy 667 GGTGTTATGCAAGGCGAGATTTGTCAGAGTTGCCGATCTGATTTGATACAGAAAGA 726
Db 670 GGTGTTATGCAAGGCGAGATTTGTCAGAGTTGCCGATCTGATTTGATACAGAGTGC 729
Qy 727 AACGACGTTGATGATGAAGCTTTCGCGGTAAATTAACAGACTATATGACTGGGT 786
Db 730 AATGACGTTGATGATGAAGCTTTCGCGGTAAATTAACAGACTATATGACTGGGT 789
Qy 787 CGACGAATTAATGCAATGATTAACAGAGGTCATGTTCCCATGTCGAGGTTCCGATTA 846
Db 790 CGACGAATTAATGCAATGATTAACAGAGGTCATGTTCCCATGTTCCGATTA 849
Qy 847 GAAATGCGCGCAAGAAATTAAGAAATGCGACCTTCAAGCGAATCTGTTTCAATCATC 906
Db 850 GAAATGCGCGCTCACTAATTAAGAAATGCGCGCTCAAGCGAATCTGTTTCAATCATC 909
Qy 907 ATGATAGCAATGATGAGGCTTGAAGGCTTTCATCTGCAAACTTATTAAGCAA 966
Db 910 ATGATAGCAATGATGAGGCTTGAAGGCTTTCATCTGCAAACTTATTAAGCAA 969
Qy 967 GCATTCAGTCAAGTGCAGAAATTCATACAAATCTCTGGGGGCGAGCGGTGAATGGGGCC 1026
Db 970 GCATTCAGTGCAGTGCAGAAATTCATACAAATCTCTGGGGGCGAGCGGTGAATGGGGCT 1029
Qy 1027 TACACGACAGATTCAGAAATGTCGATGACTATGTAAGAAATGATATGACGATTTCTT 1086
Db 1030 TACACGACAGATTCAGAAATGTCGATGACTATGTCGCAAAATGATATGACGATTTCTT 1089
Qy 1087 TTGGCGGCTGGGATTAAGGCGCGAAGCGGCTACCATGCTGCACTGCTGATACGGCTAAA 1146
Db 1090 TTGGCTCCCGGAATTAAGGCGCGAAGCGGCTACCATGCTGCACTGCTGATACGGCTAAA 1149
Qy 1147 AACGCATTAACAGTGGCGCAACGAAACCTTGCTCAAGCTTCCGTTCTATGACAT 1206

Db 1150 AATGCAATATACAGTCGAGCTACGAAACCTCCGCCAAGCTTGGCTTATTCGGAC 1209
Qy 1207 AATATTAACCAAGCTTGCACAGTTCTCTTCCGCGCCCAACAAAGATGGCGAATCAAG 1266
Db 1210 AATATCAACATATGCGACAGTTCTCTTCACTGAGCCGCAAGATGACGATCAAA 1269
Qy 1267 CCTGATGTCATGGCGCAGGGAATATCTTATTCAGCAAGATCTTCTTGAACCCGAT 1326
Db 1270 CCGAATGTCATGGCAACCGGGAAGTTCACTACAGCAAGATCTTCTTGAACCCGAT 1329
Qy 1327 TCCTCTTCTGCGCAATCAATGACAGCAATATGCTATATGAGTGAAGCTCAATGCA 1386
Db 1330 TCCTCTTCTGCGCAACATGACAGTAAATATGATACATGAGTGAAGCTCAATGCT 1389
Qy 1387 ACACCAATGTTGCGGGAATGTTGCAACAGCTCCGAGCAATTTTGAATAAGAGA 1446
Db 1390 ACACCAATGTTGCTGGAACAGTGGCAAGCTTCGAGCAATTTTGAATAAGAGGC 1449
Qy 1447 ATGACTCTTAAGCTTCCCTATGGAAGCAAGCTTATGAGTCTGCTGATGTTGA 1506
Db 1450 ATGACACCAAGCTTCTCTATTAAGCGGCACTGATGCGGTGCAAGTCAATCGGC 1509
Qy 1507 TTGGCTTATCCGAACGAAACCAAGAGTGGGCGCAAGTGAACCTGATTAATGTTGAAC 1566
Db 1510 CTGGCTTATCCGAACGATACCAAGAGTGGGCAAGTGAATGATTAATCCGTAAC 1569
Qy 1567 GTTGCCTATGTAACCAATCAGTGCCTATCACTAGCCAAAGCAATATACCTTT 1626
Db 1570 GTTGCCTATGTAACCAATCAGTGCCTATCACTAGCCAAAGCAATATGCTGTT 1629
Qy 1627 ACTGCAACGCGCGGCAAGCAATGGAATCTCCCTGATATGTCGAGTCCCTGCAAGC 1686
Db 1630 ACTGCAACGCGCGGCAAGCTTGAATCTCCCTGATATGTCGAGTCCCTGCAAGC 1689
Qy 1687 ACTACTGCTTCTGTAACCTGTCATGATGATTTGATTTGATTCAGACCAACAGCA 1746
Db 1690 ACAAGCTGCTCCGTAACGCTTGCATGATGATGATCTTGTCTTACCGCTCAAAATGCG 1749
Qy 1747 ACAAGATATGTCGGGAATGATCTTCAAGCACTTTGACAAATTAAGTGGAGCCGCAAT 1806
Db 1750 ACAAGATATGTCGGGAATGATCTTCAAGCACTTTGACAAATTAAGTGGAGCCGCAAT 1809
Qy 1807 AACGTGAAGAAATGATTTATTAATGCGCCCAAGTGAACATATCAATGAGGTGCA 1866
Db 1810 AACGTGAAGAAATGATTTATTAATGACCAACAAAGGGAAGCTATCAATGAGGTACAG 1869
Qy 1867 GCATATATGTCGCGGTGGAACCAAACTTCTGTTGGCAATTTGAACCTTA 1920
Db 1870 GCTTATTAAGTACCGGTGGAACCAACCTTCTGTTGGCAATTTGAATTTA 1923

RESULT 6
US-09-920-954-7
; Sequence 7, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGETAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-09-920-954-7

Query Match 76.9%; Score 1477.2; DB 3; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Qy 7 AAGAAAGAGTGTATCTGTTTATGAGTGGAGGATCTGTGCACTGTGCAATTA 66
Db 10 AAGAAAGAGTGTATCTGTTTATGAGTGGAGGATCTGTGCACTGTGCAATTA 69
Qy 67 AACATCCCTCGAGTGTATGCAAGCACTTTGATCTGATTTTAAAGAAATCAACA 126
Db 70 AGTAATCCATCTGCAAGTGTGCAAGAAATTTGATCTGATTTTAAAGAAATCAACA 129
Qy 127 ACACCGATGTCAGTGTCTTCTCAACAGCAACAAAGTGGCGCTGCATTTCTGTG 186
Db 130 ACACCGATGTCAGTGTCTTCTCAACAGCGGCGAGACTGTGCTGCTCTTCTGTG 189
Qy 187 GAGTCGAAAGATGTAACCTTCTTAAGGATGCTTAAGAACTTGAACAGTACCGGA 246
Db 190 GAGTCGAAAGATGTAACCTTCTTAAGGATGCTTAAGAACTTGAACAGTACCGGA 249
Qy 247 AATTAATTAATCTCATATGTCATATTCATATGAGCCCAATTTTGAAGAAACAAAG 306
Db 250 AATTAATTAATCTCATATATTCATATTCATATGAGCCCAATTTTGAAGAAACAAAG 309
Qy 307 CTAGAGCACTGAGCAAGAAATTTCTGACATATCTCGATTAATGATATTTGCGAG 366
Db 310 CTAGAGCAAGAGGCGCAAGAAATTTCTGACATATTCATATGATTAATTTGCGAG 369
Qy 367 TATGAGGAGGATGTCAGTCAAGCAAGTCCGCTCATTTGAACAGTGAATCACTGAGCCA 426
Db 370 TATGAGGAGGATGTCAGTCAAGCAAGTCCGCTCATTTGAACAGTGAATCACTGAGCCA 429
Qy 427 TACTGCGCAATTAACAAATATAGATCCCGACTTTTCACAAAGGCGATCGACGTGTG 486
Db 430 TATTTGCGCAATTAACAAATATAGATCCCGACTTTTCACAAAGGCGATCGACGTGTG 489
Qy 487 AAAGGTTGCGCTGTGATCGAAGCAACAAATTAAGAGTCAATTAAGAGCATCGAG 546
Db 490 AAAGGTTGCGCTGTGATCGAAGCAACAAATTAAGAGTCAATTAAGAGCATCGAG 549
Qy 547 GAATTCGCTCAGTATGACCAAGCAATGACGCTTATTAATTAAGGCAAGCTTGAAT 606
Db 550 GAATTCGCAATTTGCAATTAAGCAATGATGCTATTAATTAAGGCAAGCTTGAAT 609
Qy 607 AAGGTGATGATGATGTCGAGAGGATTTTGAAGGAGTGTGCAAGGAGTGTGCAAGGAGCTAC 666
Db 610 AAGGTGATGATGATGTCGAGAGGATTTTGAAGGAGTGTGCAAGGAGTGTGCAAGGAGCTAC 669
Qy 667 GGTTTGATGCAAGGCGCAATTTGTCAGTGGCCGATTACTGATTTGATTAAGAGGA 726
Db 670 GGTTTGATGCAAGGCGCAATTTGTCAGTGGCCGATTACTGATTTGATTAAGAGGA 729
Qy 727 AACGACGATGATGCAAGGCGCTTCCGCGTAAATTAACGCACTTAATGACCTGGGT 786
Db 730 AATGACGATGATGCAAGGCGCTTCCGCGTAAATTAACGCACTTAATGACCTGGGT 789
Qy 787 CGGAGCAATTAATGCAATGATACGAACGTCATGTCATGTCAGTGTGCGATTTA 846
Db 790 CGGAGCAATTAATGCAATGATACGAATGTCATGTCATGTCAGTGTGCGATTTA 849

847 GGAATGCGGCAAGAAATGAAGAAATGCGACCTCAAGCGAATCTGGTTTTCATCCATC 906
850 GGAACCGGCTCACTAATTAAGAAATGGGCTCGAGCGAATCTGATCTTCCAACTATC 909
907 ATGATAGCAAGTGTGGGCTTGGAGGCTTGCCTCCATCTGCAAACTTATTCAGCCA 966
910 ATGATAGCGGTGGGGACTTGGAGGACTACCTTGGAACTTCCAAACCTTATTCAGCCA 969
967 GCATTCAGTGCAGGTGGCCAGAAATTCATCAAACTCTGGGGGGCGAGCGGTGAATGGGCT 1026
970 GCATTCAGTGCAGGTGGCCAGAAATTCATCAAACTCTGGGGGGCGAGCGGTGAATGGGCT 1029
1027 TACACACAGATTCAGAAATGTGATGATCTATGTAAGGAAATGATATGACATCTT 1086
1030 TACACACAGATTCAGAAATGTGATGATCTATGTAAGGAAATGATATGACATCTT 1089
1087 TTCCGGGCTGGGAATGAAGAGCCGAAACGGCGGTACCATCATGTCACCTGGTAA 1146
1090 TTCCGGGCTGGGAATGAAGAGCCGAAACGGCGGTACCATCATGTCACCTGGTAA 1149
1147 AACGCAATACAGTGGGCGCAACCGAAACCTGCGTCCAGCTTGGTTCTATGCAAT 1206
1150 AATGCAATACAGTGGGCGCAACCGAAACCTGCGTCCAGCTTGGTTCTATGCAAT 1209
1207 AATATTAACAGTGGCAAGTTCTCTCCGGTGGCCGCAAAAGATGGGCGAATCAG 1266
1210 AATATTAACAGTGGCAAGTTCTCTCCGGTGGCCGCAAAAGATGGGCGAATCAG 1269
1267 CCTGATGTCATGGCGCCAGGGAATATATTTATGAGCAAGATCTTCTTGGCAACCAT 1326
1270 CCGATATGTCATGGCGCCAGGGAATATATTTATGAGCAAGATCTTCTTGGCAACCAT 1329
1327 TCCTCTCTTGGGCGCAATCATACAGCAAAATATGCTATATGGTGGAAACGTCCATGCA 1386
1330 TCCTCTCTTGGGCGCAACCATATACAGTAATATGATATGATGGTGGAAACGTCCATGCT 1389
1387 ACACGATTTGTTGGGGAATGTTGCAACGCTCCGAGCAATTTGTGAAATAAGAGA 1446
1390 ACACGATTTGTTGGGGAATGTTGCAACGCTCCGAGCAATTTGTGAAATAAGAGA 1449
1447 ATCACTCTTAAGCCTTCCCTATGGAAGCAAGCTTTGATGCAAGTGTGATGTTG 1506
1450 ATCACTCTTAAGCCTTCCCTATGGAAGCAAGCTTTGATGCAAGTGTGATGTTG 1509
1507 TTGGTTATCCGGAAGGAAACCAAGATGGGGCGAGTGAACCTCGATTAATCGTTGAAC 1566
1510 CTGGCTACCCGAAAGGTAACCAAGATGGGGCGAGTGAACCTCGATTAATCGTTGAAC 1569
1567 GTTGCCTATGTAAGCAATCCAGTCCCTATCACTAGCAAAAGCGACATATACCTTT 1626
1570 GTTGCCTATGTAAGCAATCCAGTCCCTATCACTAGCAAAAGCGACATATACCTTT 1629
1637 ACTGCAAGCGGCGGAGCAATGGAATCTCCCTGATGATGATGATGATGATGATGATG 1686
1630 ACTGCAAGCGGCGGAGCAATGGAATCTCCCTGATGATGATGATGATGATGATGATG 1689
1687 ACTACTGCTTCTGTAACCTGCTGTAATGATTTGATTTGATTTGATTTGATTTGATTTG 1746
1690 ACTACTGCTTCTGTAACCTGCTGTAATGATTTGATTTGATTTGATTTGATTTGATTTG 1749
1747 ACAAGATATGTGCGGAATGACTTCTGAGCAACATTTGCAATTAATCTGGAATGGCCGCAAT 1806
1750 ACAAGATATGTGCGGAATGACTTCTGAGCAACATTTGCAATTAATCTGGAATGGCCGCAAT 1809
1807 AACGTAAGAAATGTAATTTATTAATTCGCCCAAGATGGAACATATACATTTGAGGTGCA 1866
1810 AACGTAAGAAATGTAATTTATTAATTCACCAACAGCGGAGCGTATATCAATTTGAGGTGCA 1869
1867 GCATTAATGTGCGGCTGGAACCAAAATTTCTGTTGGCAATGTTGAACTGA 1920
1870 GCATTAATGTGCGGCTGGAACCAAAATTTCTGTTGGCAATGTTGAACTGA 1923

RESULT 7
US-08-873-479-41
; Sequence 41, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Slome, Alan
; APPLICANT: Lynne, Christlanson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 58917010 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251,000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3003 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-873-479-41

Query Match 59.2%; Score 1136.8; DB 2; Length 3003;
Best Local Similarity 74.9%; Pred. No. 0;
Matches 1436; Conservative 0; Mismatches 477; Indels 3; Gaps 1;

5 GAAAGAAAGAGGTGTTTATCTGTTTATCAAGCTGACGATTCGTGCACTGTGCAT 64
659 GATGGAAGAGGTTTTTTTATCCGTTTATCAAGTGTGCACTATGTTCTTCTGTTGCTT 918
65 TAAACAAATCCCTCGGCTGGTGAATGCAAGACTTTGATCTGGAATTTTAAAGAAATTCAA 124
919 TAAACAAATCCCTCGGCTGGTGAATGCAAGACTTTGATCTGGAATTTTAAAGAAATTCAA 978
125 CAACAAACCAATGTCAGTGTCTTCTCCAAACAGCAACAAACAGTGGCGGTGCAATTTCTGG 184
979 CACTTAAGCTAGAGAGAGGCTGCAACCAAGAAAGAAAGAGGAAAGAGGAAAGAGGAAAG 1038
185 TGAGTCTGAAATAGTGAATCTTCTTAAAGAAATGCTTAAAGAAATCTGAAACAGTACCGG 244
1039 TAAACCTGAAATAGTGAATCTTCTTAAAGAAATGCTTAAAGAAATCTGAAACAGTACCGG 1098
245 CAATTAATTAATCTCATATGTTGCAATTAATTAATGAGGAGGAGGAGGAGGAGGAGGAGG 304
1099 CGGATTAACAGCTATATATGTTGCAATTAATTAATGAGGAGGAGGAGGAGGAGGAGGAGG 1158
305 AGCTAGAGCAACCTGAGGCAAGATTTCTGACATCACTCCCTGATATGCAATATTTCTGG 364
1159 AACTAGAGAGAGCGGAGGAGGAAATTTCTGATTAATCAACAGATTAATGCTTATATTTCTG 1218
365 AGTATGAGGAGGAGTGTCACTGCAAAAGTCCGCTCAATTTGAAACAGTGAATCAAGTGAAGC 424

Db 1219 AATATGATGGGGATGTAAGAGCCGTAACCTPAACGAATTCGGCATTTTGAATCGGTTGAAC 1278
QY 425 CATCTTGGCCGAATACAAATATAGATCCCAAGCTTTTTCACAAAAGGCGCATCGACCTGG 484
Db 1279 CATATTACCTTATATATAAATAGACCCGCAATTAATTTTTCAGAGAGCTTCCTGAATTAG 1338
QY 485 TGAAGCGTTGGCGCTTGATACGAGACGAAACAATAAGAGTGAATTAAGAGCATG 544
Db 1339 TAGAAACAGTAGCTTTAGATTAAGAAAGCAAGAGTAAAGAGTACCTTTAAGAGATTGG 1398
QY 545 AGGAATTCCTCGATACGTAAGCAAGCAATGACGTCAATATATTAACGCAAAAGCTGAAT 604
Db 1399 AACAAATGCCCAATACGCAACAATATGATGTAATTATACGTAAACCCAAAGCTGAAT 1458
QY 605 ATTAAGGTATGAAATGATGTGGCCAGAGATTTGTCAAAAGCGATGTGGCAGAGACGCT 664
Db 1459 ACGAATTTTGAATGACGTGGCCCGGTGCAATGTGAAAGCAAGACGTGCAAAAATACCT 1518
QY 665 ACGGTTTATGGAACAAGGCGAGATTGTGCGAGTTGCGATACCTGATGATTTGAATGAGAA 724
Db 1519 TTGGCTTAATGACAAAGACAGATTTGTAGCACTGTCTGATACCTGGCTTGAATACAGAA 1578
QY 725 GAACACAGATTCGATGACGTAAGACCTTCGCGGTAAATTAACGACATTAATGACCTGG 784
Db 1579 GAAATGACAGTTTCGATGACGTAAGACATTCGCGGTAAAGATTACCGACATTAATGACCTGG 1638
QY 785 GTCCGACGAAATATGGAATGATACGAAGGTGATGTACCCATGTGGCAGGTTCCGTAT 844
Db 1639 GCAAGAACGATTAACGCAATGATCCAAATGACATGAAACCCATGTTGCTGATCTGTGT 1698
QY 845 TAGGAATGGCGCAACGAATTAAGAAATGGCACTCAAGCGAATCTGTTTTTCAATCCA 904
Db 1699 TAGGAAT---GCTAACAATTAAGGATGACACCGAAGCGCAATCAATGCTTTCAATCTA 1755
QY 905 TCATGATATGACAGTGTGGCTTGAAGGCTTGCCTTCAATCTGCAAACTTAATTCAGCC 964
Db 1756 TTATGATATGATGTGAGAGGCTGGAGGACATCTGCTATATCAACAACTTAATTCAGTC 1815
QY 965 AAGCATTCAGTGCAGTGCAGAAATTCATACAACTCTGGGGGGGAGCGGTGATGAGCG 1024
Db 1816 AAGCATATGATGTGAGAGGAGCAATTCATACGAATTCATGAGGGGCGCTCAAGTAAACG 1875
QY 1025 CCTACACGACAGATTCAGAAATGTGATGACATATGTAAGAAATATGATGACGATTC 1084
Db 1876 CCTATACGACAGCTTCGAAATGTGATGATGTAAGAAATATGATGACGATTC 1935
QY 1085 TTTTCCGCGCTGGAAATGAAGGCGGAGTACATAGTGCACCTGAGTACGGCTA 1144
Db 1936 TTTTTCGGGCGGAAATGAGGAGCAAGTACGCTACATGATGACCAAGAAACGCA 1995
QY 1145 AAAAGCCATTAACAGTGGCGCAACGAAACCTGCGTCAAGCTTCGATTCATGCG 1204
Db 1996 AAAATGCGATTACAGTTGGGGCAACGAAACCTAGCTCAAGCTTCGATTCATGCG 2055
QY 1205 ATTAATTTTAAACAGCTTGAACAGTTCTCTCCGTTGGCCCGCAAAAGATGGCCAAATCA 1264
Db 2056 ATTAATTTTAAACAGTGTGCTCAATTCCTTCAAGAGTCTCTAAGATGACGATTA 2115
QY 1265 AGCTGATATGACGCGGAGGAGCAATACATTTTATCAGACAGATCTTCTTTCGACCG 1324
Db 2116 AGCCGAGCTCATGAGCAACAGTACGTATATTTCTCTGCTAGATCATTTAGCTCCAG 2175
QY 1335 ATTCTCTCTTGGGCGAAATCATGACGAAATATGCTATATATGAGTGAACGTCATG 1384
Db 2176 ATTCTCTCATTTGGGCAAAACATGATGATTAATGCTATACATGAGTGTATCTTATG 2235
QY 1385 CAACACCAATTTGTTGGGGGAATGTTGCAAGTCCGTACAGATTTTGGAAAAATAGAG 1444
Db 2236 CTACTCCAAATTTGAGAGGTAAATGTTGCAATTAAGGAGCAATTTTGGAAAAATAGAG 2295
QY 1445 GAATCACTCTTAAGCTTCCCTAATTAAGACAGCTTGAATGAGGTGCTGATGATG 1504
Db 2296 GGGTAACTCTTAAGCTTCCCTTAAAGCTGCTTTAATTTGACAGGTGCTGCGAGATTTG 2355

QY 1505 GATTGGGTTATGCGAACGGAACCAAGATGAGGCGGAGTGAACCTGATTAATCGTTGA 1564
Db 2356 GACTTGGCTTTCCAAATGATTAACAAAGATGGGAAAGATTAACGTTAGATTAATCCCTTA 2415
QY 1565 ACGTGGCTATGTGAACGAATCAGTGCCTTATCAACTAGCCAAAAAGGACATATACCT 1624
Db 2416 ATGTGCAATTTGTGAATGAAGAGAGCCCTTTTCAACAGTCAAAAAGCAATATTCGT 2475
QY 1625 TTATGCAACGCGGCGCAAGCCATTGAAATCTCCCTGATGTGCGATGCCCCCTGCA 1684
Db 2476 TTACGCGCTCAAGCTGTAACCTTAAATATCACTTGTGTGTCAGATGACAGTGA 2535
QY 1685 GCACTACTGCTCTGTAAACCCGTGTCATATGATTTGGTCACTTAAGCAACCAAG 1744
Db 2536 GCACGACGCGCATCACTTAACCTTAATGTAATGATTTAGATTAATGATCACTGCAACCAAG 2595
QY 1745 GAACAAAGATATGTCGGGAATGACTTCTGACGACATTTGACAAATACGGAATGCGCGCA 1804
Db 2596 GAACATTAATGCTCGAAATGACTTTACAGCACCGTATGATTAACATTTGGATGGCAAG 2655
QY 1805 ATTAAGTAAATATGATTTATTAATTCGCCCCAAAGTGAACATATACATTGAGTGC 1864
Db 2656 ACAAGTGAATATGTTTATCAATGCTCTCAAGCGGAACGTATACATGCTGAAGTGC 2715
QY 1865 AAGCATATATGTCGCGGTGGAACCAAAACTTCTCGTTGGCAATTTGAACCTAA 1920
Db 2716 AAGCTTAACATGATCAAGTAAAGTCCGCAAACTTTTCTTACGATTTGATCAATTA 2771

RESULT 8
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHRIFFLINGER, F.
; APPLICANT: PALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0239
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: BP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELE: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid


```
Db      609  ATTAAAAACATTGAAGATTCAACAGAAAGTGGCTAAAAATGTTTGAAGCAACGCTTC 668
Qy      372  GGGGAGTGTTCAGTCAAAAGTCGGCTCCATTGAACAAGTGAATCAGTGAAGCCACTT 431
Db      669  TGAAGAAACAGAGGAAAAAAGACGGCTTCACTTAAACAGAAAAAAGAGGAAAAACGTTT 728
Qy      432  GCCGAATACAAATATGATCCCACTTTTCACAAAAGCCGATGACGCTGTGAAGC 491
Db      729  AGCAGAAAGAACAGACGTCACGTCAGCTGCTTAAAAAGCAAGAAACAAAGCCGACG 788
Qy      492  GTTGGCGCTGTATAGAAAGCAGACAAATTAAGAGTGAATTAAGAGGCATGAGAAAT 551
Db      789  GCAGACCAAGCCCGACGACCAAAAGCCGACAGACAGCAAGCAAAACAAAGCAGC 848
Qy      552  CGCTCAGTACGTAGCA 567
Db      849  CAATGAGGACAGCAGCA 864
```

```
RESULT 14
US-09-949-016-815
; Sequence 815, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 815
; LENGTH: 2947
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-815
```

```
Query Match      2.1%; Score 39.6; DB 3; Length 2947;
Best Local Similarity 44.1%; Pred. No. 0.15;
Matches 165; Conservative 0; Mismatches 209; Indels 0; Gaps 0;
```

```
Qy      151  AAACGCGCAACAAACGATGCGGCTGCTATTTCTGTGAGTCTGAATAATGTAACTTCTT 210
Db      443  AAAGAAATACAAAGATTAAGCTGAAAAACAAAGAAACCTGAAAGAAATACAAACT 502
Qy      211  AAAGATTTGCTTAAGAACTTGAACAGTACCGCAATTAATTAATCTCATTTGTCCA 270
Db      503  AAAGTTACACAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAACT 562
Qy      271  TTCAATGCCCCCATTTTGAAGAAACAAACAGAACTAGAGCAACTGAGCAAAAGATT 330
Db      563  GAAAAAGAAAGCAACTCAAGAGAAAAAATTGAAAAAAGAAAAACAGAAAAAGAAACA 622
Qy      331  CTGCACTACATCCCTGATTTATGATATATGTCAGATAGAGGGAGATGTTCAATCAAA 390
Db      623  GTGGGAAAGAAACAGAAAGAAAGCTAAGACTGCAAGAAAGAGTGAAGAAAAAGCTAAAAG 682
Qy      391  GTCCGCTCCATTGAACAGTGAATCAGTGAAGCCATCTTGCCGAAATTAACAAATAGAT 450
Db      683  GAAGTGAAGAGTGAAGAAAGAGAAAGTGAAGCAACGCTGCAAAAGTAAAGAAAGTA 742
Qy      451  CCCAGCTTTTCAAAAGGCGCATGACGCTGTGAAGCGTTGCGCTTGATACGAAG 510
Db      743  CAGAAACACCATCAAAACCCAAAGAAAAAGAGAGCAAAAGAGAAAGCAGCTGTCTAAAG 802
```

```
Qy      511  CAGAACATTAAGA 524
Db      803  CATGAACGAAAGA 816
```

```
RESULT 15
US-09-824-893A-50
; Sequence 50, Application US/09824893A
; Patent No. 6902922
; GENERAL INFORMATION:
; APPLICANT: Maxygen Inc.
; APPLICANT: No. 6902922ozymes A/S
; APPLICANT: Nees, Jon
; APPLICANT: Giver, Lorraine
; APPLICANT: Cherly, Joel
; APPLICANT: Borchert, Torben
; APPLICANT: Stemmer, Willem
; APPLICANT: Minshull, Jeremy
; TITLE OF INVENTION: Subtilisin Variants
; FILE REFERENCE: 10181,200-US
; CURRENT APPLICATION NUMBER: US/09/824,893A
; CURRENT FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 270
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-824-893A-50
```

```
Query Match      2.0%; Score 39; DB 3; Length 522;
Best Local Similarity 68.4%; Pred. No. 0.079;
Matches 54; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
```

```
Qy      1336  TGGGCAATCATGACAGAAATATGCTATATGGTGAAGCTCCAGCAACCCGATT 1395
Db      440  TCGACATATCGGGCAGAGATATGCTTCAATGAGCGGACATCAATGCAACCGCAT 499
Qy      1396  GTTCCGGGAATGTTGCAC 1414
Db      500  GTTCAAGCGCTTGACGAC 518
```

```
Search completed: April 8, 2006, 00:49:20
Job time : 592.04 secs
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 01:34:45 ; Search time 1378.42 Seconds
(without alignments)
11518.383 Million cell updates/sec

Title: US-10-784-870-3

Perfect score: 1920

Sequence: 1 atgagagaagaagaagtgctt.....cgcttgcaattgtgaactaa 1920

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413469005 residues

Total number of hits satisfying chosen parameters: 19567084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1920	100.0	1920	3	US-09-920-954-3
2	1920	100.0	1920	7	US-10-784-870-3
3	1477.2	76.9	1923	3	US-09-920-954-5
4	1477.2	76.9	1923	3	US-09-920-954-7
5	1477.2	76.9	1923	7	US-10-456-479-3
6	1477.2	76.9	1923	7	US-10-784-870-5
7	1477.2	76.9	1923	7	US-10-784-870-7
8	1477.2	76.9	1923	7	US-10-820-712A-2
9	1477.2	76.9	1923	8	US-10-820-712A-2
10	1295.6	67.5	1302	8	US-10-820-712A-11
11	1295.6	67.5	1302	8	US-10-820-712A-12
12	1002.6	52.2	1305	6	US-10-385-662-1
13	999.6	52.1	1302	8	US-10-820-712A-13
14	999.6	52.1	1302	8	US-10-820-712A-14
15	823.6	42.9	1302	8	US-10-820-712A-21
16	823.6	42.9	1302	8	US-10-820-712A-22
17	747.6	38.9	1299	8	US-10-820-712A-17
18	747.6	38.9	1299	8	US-10-820-712A-18
19	738	38.4	1299	8	US-10-820-712A-19
20	734.8	38.4	1299	8	US-10-820-712A-20
21	734.8	38.3	1299	8	US-10-820-712A-15
22	734.8	38.3	1299	9	US-10-820-712A-16
23	40	2.1	522	3	US-09-824-893A-56

24	40	2.1	522	3	US-09-824-893A-70	Sequence 70, Appl
25	40	2.1	522	8	US-10-736-997-56	Sequence 56, Appl
26	40	2.1	522	8	US-10-736-997-70	Sequence 70, Appl
27	40	2.1	8911	3	US-09-070-927A-651	Sequence 651, Appl
28	39.6	2.1	2382	5	US-10-084-817-312	Sequence 312, Appl
29	39.6	2.1	2947	9	US-10-287-436A-213	Sequence 213, Appl
30	39.6	2.1	4153	8	US-10-723-860-5175	Sequence 5175, Appl
31	39	2.0	522	3	US-09-824-893A-50	Sequence 50, Appl
32	39	2.0	522	8	US-10-736-997-50	Sequence 50, Appl
33	38.6	2.0	1192	7	US-10-282-122A-35745	Sequence 35745, A
34	38.4	2.0	522	3	US-09-824-893A-37	Sequence 37, Appl
35	38.4	2.0	522	3	US-09-824-893A-43	Sequence 43, Appl
36	38.4	2.0	522	3	US-09-824-893A-47	Sequence 47, Appl
37	38.4	2.0	522	3	US-09-824-893A-60	Sequence 60, Appl
38	38.4	2.0	522	3	US-09-824-893A-62	Sequence 62, Appl
39	38.4	2.0	522	3	US-09-824-893A-66	Sequence 66, Appl
40	38.4	2.0	522	3	US-09-824-893A-76	Sequence 76, Appl
41	38.4	2.0	522	3	US-09-824-893A-79	Sequence 79, Appl
42	38.4	2.0	522	3	US-09-824-893A-80	Sequence 80, Appl
43	38.4	2.0	522	3	US-09-824-893A-83	Sequence 83, Appl
44	38.4	2.0	522	3	US-09-824-893A-101	Sequence 101, Appl
45	38.4	2.0	522	3	US-09-824-893A-112	Sequence 112, Appl

ALIGNMENTS

```
RESULT 1
US-09-920-954-3
Sequence 3, Application US/09920954
Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIRATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 1920
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURES:
NAME/KEY: CDS
LOCATION: (1)..(1920)
US-09-920-954-3

Query Match 100.0%; Score 1920; DB 3; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAAAAGAGAGAGTGTGTTTATCTGTTTATGCTGACGCACTTCTGCACTGTT 60
DB 1 ATGAGAAAAGAGAGAGTGTGTTTATCTGTTTATGCTGACGCACTTCTGCACTGTT 60
QY 61 GCATTAACAATCCCTCGGCTGCTGATGCAAGCACTTTGATGATTTTAAAGAAATT 120
DB 61 GCATTAACAATCCCTCGGCTGCTGATGCAAGCACTTTGATGATTTTAAAGAAATT 120
```

121 CAAACAACACCGATGTGAGTGGTTCTCCAAACAGCGACAAACAGGTGCGTGCATT 180
121 CAAACAACACCGATGTGAGTGGTTCTCCAAACAGCGACAAACAGGTGCGTGCATT 180
181 CTGGTGGAGTCTGAAAATGTGAACTCTTAAAGATTGCTAAAGAACTTGAACAAGTA 240
181 CTGGTGGAGTCTGAAAATGTGAACTCTTAAAGATTGCTAAAGAACTTGAACAAGTA 240
241 CCGGCAAAATTAATTAATCTCAATATGTCGAATTAATGAGCCCAATTTAAGAAACAAA 300
241 CCGGCAAAATTAATTAATCTCAATATGTCGAATTAATGAGCCCAATTTAAGAAACAAA 300
301 CAGAACTAGAGCAACTGAGCAAAAGATTCTGACTACATCCCTGATTAATGATATAT 360
301 CAGAACTAGAGCAACTGAGCAAAAGATTCTGACTACATCCCTGATTAATGATATAT 360
361 GTGAGTATGAGGGGAGTGTCAAGTCAAAAGTCCGCTCATTTGAACACGTGGAATCAGTG 420
361 GTGAGTATGAGGGGAGTGTCAAGTCAAAAGTCCGCTCATTTGAACACGTGGAATCAGTG 420
421 GAGCCATCTTGCCGAATTAATTAATGATCCCGAGCTTTGACAAAGGCGCATCGACG 480
421 GAGCCATCTTGCCGAATTAATTAATGATCCCGAGCTTTGACAAAGGCGCATCGACG 480
481 CTGGTGAAGCGTTGGCGCTTGATACGAAGCAGAAACAATAAAGAGTGAATTAAGAGGC 540
481 CTGGTGAAGCGTTGGCGCTTGATACGAAGCAGAAACAATAAAGAGTGAATTAAGAGGC 540
541 ATGAGGAATTCGCTCAGTACGTAAGCAATGACGTCCATTAATTAATGAGCAAAAGCT 600
541 ATGAGGAATTCGCTCAGTACGTAAGCAATGACGTCCATTAATTAATGAGCAAAAGCT 600
601 GAAATTAAGGTATGATGATGATGAGCAGAGGATTTGCAAGCGGAGTGGCAGACAGC 660
601 GAAATTAAGGTATGATGATGATGAGCAGAGGATTTGCAAGCGGAGTGGCAGACAGC 660
661 AGCTACGCTTTGTATGACAAAGCCAGATGTGTGCGAGTTGCCGATCTGATGATGATCA 720
661 AGCTACGCTTTGTATGACAAAGCCAGATGTGTGCGAGTTGCCGATCTGATGATGATCA 720
721 GGAAGAAACGACGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
721 GGAAGAAACGACGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
781 GGAAGAAACGACGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
781 GGAAGAAACGACGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
841 GTATTAGAAATGAGCGCAACGAATTAAGGAATGAGCACTCAAGCGAATCTGTTTTCGA 900
841 GTATTAGAAATGAGCGCAACGAATTAAGGAATGAGCACTCAAGCGAATCTGTTTTCGA 900
901 TCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
901 TCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
961 AGCCAAGCATTCAGTGAAGTGGTCCAGAAATTCATACAACTCTGAGGGGCGAGCGGTGAT 1020
961 AGCCAAGCATTCAGTGAAGTGGTCCAGAAATTCATACAACTCTGAGGGGCGAGCGGTGAT 1020
1021 GGGGCTTACAGCAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1080
1021 GGGGCTTACAGCAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1080
1081 ATTCTTTTGGGGCTGGGAATGAAAGCGCGAATCGCGGTACATGATGATGATGATGATG 1140
1081 ATTCTTTTGGGGCTGGGAATGAAAGCGCGAATCGCGGTACATGATGATGATGATGATG 1140
1141 GCTAAATACGCAATTAAGTGGCGCAACCGAAACCTGCGTCAAGGCTTGGGTTCCAT 1200
1141 GCTAAATACGCAATTAAGTGGCGCAACCGAAACCTGCGTCAAGGCTTGGGTTCCAT 1200
1201 GCAATTAATTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
1201 GCAATTAATTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1260

1201 GCAATTAATTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1321 CCGAATTCCTCTTCTGAGCGAATCATGACGCAAAATGATGATGATGATGATGATGATGATG 1380
1321 CCGAATTCCTCTTCTGAGCGAATCATGACGCAAAATGATGATGATGATGATGATGATGATG 1380
1381 ATGCAACACCGAATTTGTTGGCGGAAATGTTGCAAGCTCCGTGAGCAATTTGTTGAATA 1440
1381 ATGCAACACCGAATTTGTTGGCGGAAATGTTGCAAGCTCCGTGAGCAATTTGTTGAATA 1440
1441 AGAGGAATCACTCTTAAGCTTCCCTTAATGAAAGCAGCTTGAATGATGATGATGATGAT 1500
1441 AGAGGAATCACTCTTAAGCTTCCCTTAATGAAAGCAGCTTGAATGATGATGATGATGAT 1500
1501 GTTGAATGGGTTATCCGAACGGAACCAAGATGAGGCGAGTGAACCTGATTAATTCG 1560
1501 GTTGAATGGGTTATCCGAACGGAACCAAGATGAGGCGAGTGAACCTGATTAATTCG 1560
1561 TTGAACCTTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1561 TTGAACCTTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1621 ACCTTACTGCAACGCGGCGGAGCCATGAAATCTCCGTGATGATGATGATGATGATGAT 1680
1621 ACCTTACTGCAACGCGGCGGAGCCATGAAATCTCCGTGATGATGATGATGATGATGAT 1680
1681 GCAAGCACTACTGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1681 GCAAGCACTACTGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1741 AACGGAACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
1741 AACGGAACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
1801 CGCAATTAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
1801 CGCAATTAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
1861 GTGCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
1861 GTGCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920

RESULT 2
US-10-784-870-3
Sequence 3, Application US/10784870
Publication No. US20040142837A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0BCT
CURRENT APPLICATION NUMBER: US/10/784, 870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509, 814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0

/ SEQ ID NO 3
/ LENGTH: 1920
/ TYPE: DNA
/ ORGANISM: Bacillus sp.
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1) .. (1920)
US-10-784-870-3

Query Match 100.0%; Score 1920; DB 7; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGAAAGAAAGAGGTTTATCTGTTTATCAGCTCCAGCCGATCTGTGACTGTT 60
DB 1 ATGAGAAAGAAAGAGGTTTATCTGTTTATCAGCTCCAGCCGATCTGTGACTGTT 60
QY 61 GCATTAAACATCCCTCGCTGGTGAAGCAAGGACTTTGATCTGATTTTAAAGAAATT 120
DB 61 GCATTAAACATCCCTCGCTGGTGAAGCAAGGACTTTGATCTGATTTTAAAGAAATT 120
QY 121 CAAACAAACACGATGTCAGTGTCTTCMAACAGCGACAAACAGGTGCGCTGCATTT 180
DB 121 CAAACAAACACGATGTCAGTGTCTTCMAACAGCGACAAACAGGTGCGCTGCATTT 180
QY 181 CTGGTGAAGTCTGAAAATGTGAACTTCTTAAAGATTCGTAAAGAACTTGAAACAGTA 240
DB 181 CTGGTGAAGTCTGAAAATGTGAACTTCTTAAAGATTCGTAAAGAACTTGAAACAGTA 240
QY 241 CCGGCAATATATTAATCTCATATATGTCATTCATTCAGTGGCCCATTTTAAAGAAACAAA 300
DB 241 CCGGCAATATATTAATCTCATATATGTCATTCATTCAGTGGCCCATTTTAAAGAAACAAA 300
QY 301 CAGAACTTAGAGCAACCTGAGCAAAAGATTCTGACTACATCCCTGATTAATGATATATT 360
DB 301 CAGAACTTAGAGCAACCTGAGCAAAAGATTCTGACTACATCCCTGATTAATGATATATT 360
QY 361 GTTCAGATAGAGGGGGATGTTCAGTCAAAAGTCCGCTCCATTTGAACAAGTGAATGATG 420
DB 361 GTTCAGATAGAGGGGGATGTTCAGTCAAAAGTCCGCTCCATTTGAACAAGTGAATGATG 420
QY 421 GAGCCATACCTGCGCAATATACAAATATAGATCCCACTTTTCAAAAAGCGCATGAGG 480
DB 421 GAGCCATACCTGCGCAATATACAAATATAGATCCCACTTTTCAAAAAGCGCATGAGG 480
QY 481 CTGGTGAAGCGTTCGCTTGAATGAGAGCAAGAACATTAAGAGTGCATTTAAGAGGC 540
DB 481 CTGGTGAAGCGTTCGCTTGAATGAGAGCAAGAACATTAAGAGTGCATTTAAGAGGC 540
QY 541 ATCGAGGAATGCTGATGATGCTAGAGCAATGACGTCTCATTTATTAACGCAAGCCT 600
DB 541 ATCGAGGAATGCTGATGATGCTAGAGCAATGACGTCTCATTTATTAACGCAAGCCT 600
QY 601 GAAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 GAAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 AGCTACGCTTTGATGAGCAAGGCGCATGTTGCGCAGTTGCGATCTGATTTGATGATA 720
DB 661 AGCTACGCTTTGATGAGCAAGGCGCATGTTGCGCAGTTGCGATCTGATTTGATGATA 720
QY 721 GGAAGAAACGACATGTCATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 721 GGAAGAAACGACATGTCATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 781 CTGGGTGCGACGAATTAATGCGAATGATGATGATGATGATGATGATGATGATGATGATG 840
DB 781 CTGGGTGCGACGAATTAATGCGAATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 841 GTATTAGGAATGCGCAACGAATTAAGGAATGCGCACTCAAGCGAATCTGCTTTTCA 900
DB 841 GTATTAGGAATGCGCAACGAATTAAGGAATGCGCACTCAAGCGAATCTGCTTTTCA 900
```

```
QY 901 TCCATCATGATATAGAGTGTGGGCTTGGAGGCTTGCCTTCCAACTCGAAACCTTATTC 960
DB 901 TCCATCATGATATAGAGTGTGGGCTTGGAGGCTTGCCTTCCAACTCGAAACCTTATTC 960
QY 961 AGCCAGCATTCAGTGCAGGTGTCAGAAATTCATACAACTCTGGGGGCGACGGTGAAT 1020
DB 961 AGCCAGCATTCAGTGCAGGTGTCAGAAATTCATACAACTCTGGGGGCGACGGTGAAT 1020
QY 1021 GGGGCTTACACACAGCAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATG 1080
DB 1021 GGGGCTTACACACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 ATTCTTTTGGCGGCTGGGAATGAAAGCCGGAACGGCGGTACCATGATGATGATGATG 1140
DB 1081 ATTCTTTTGGCGGCTGGGAATGAAAGCCGGAACGGCGGTACCATGATGATGATGATG 1140
QY 1141 GCTAAAGAAACGCAATTAACAGTTCGACCAACGAAACCTGCTCAAGCTTCGGTCTTAT 1200
DB 1141 GCTAAAGAAACGCAATTAACAGTTCGACCAACGAAACCTGCTCAAGCTTCGGTCTTAT 1200
QY 1201 GCAGATTAATATTAACCAAGTTCGACAGTTCCTTCCGCTGACCAAAAGATGGGCGA 1260
DB 1201 GCAGATTAATATTAACCAAGTTCGACAGTTCCTTCCGCTGACCAAAAGATGGGCGA 1260
QY 1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
DB 1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
QY 1321 CCCGATTCCTCTCTGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1321 CCCGATTCCTCTCTGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
QY 1381 ATGGCAACACCGATTTGTTGGGGGAATGTTGCAAGCTTCGATGAGCAATTTTGAAGAAAT 1440
DB 1381 ATGGCAACACCGATTTGTTGGGGGAATGTTGCAAGCTTCGATGAGCAATTTTGAAGAAAT 1440
QY 1441 AGAGGAATCACTCTTAAGCTTCTCTATTTGAAGCAAGCTTTGATGATGATGATGATGAT 1500
DB 1441 AGAGGAATCACTCTTAAGCTTCTCTATTTGAAGCAAGCTTTGATGATGATGATGATGAT 1500
QY 1501 GTTGAATGGGTTATCCGAACGGAACCAAGGATGGGGCCGAGTGAACCTGATTAATATG 1560
DB 1501 GTTGAATGGGTTATCCGAACGGAACCAAGGATGGGGCCGAGTGAACCTGATTAATATG 1560
QY 1561 TTGAACGTTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
DB 1561 TTGAACGTTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1621 ACCTTTTACTGCAACGGCGGCAAGCATTGAAATCTCCTGATATGATGATGATGATG 1680
DB 1621 ACCTTTTACTGCAACGGCGGCAAGCATTGAAATCTCCTGATATGATGATGATGATGATG 1680
QY 1681 GGAAGCACTACGCTTCTGATTAACCTGATGATGATGATGATGATGATGATGATGATGATG 1740
DB 1681 GGAAGCACTACGCTTCTGATTAACCTGATGATGATGATGATGATGATGATGATGATGATG 1740
QY 1741 AACGGAACAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
DB 1741 AACGGAACAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
QY 1801 CGCAATTAACGTAAGAAATGATATTTATTAATGCGCCCAAGTGAACATTAACATTTGAG 1860
DB 1801 CGCAATTAACGTAAGAAATGATATTTATTAATGCGCCCAAGTGAACATTAACATTTGAG 1860
QY 1861 GTGCAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
DB 1861 GTGCAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
```

RESULT 3
US-09-920-954-5
; Sequence 5. Application US/09920954
; Publication No. US20020064854A1

GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
PRIORITY FILING DATE: 2001-08-03
PRIORITY APPLICATION NUMBER: 09/509,814
PRIORITY FILING DATE: 2000-04-06
PRIORITY APPLICATION NUMBER: PCT/JP98/04528
PRIORITY FILING DATE: 1998-10-07
PRIORITY APPLICATION NUMBER: JP 9-274570
PRIORITY FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 5
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-09-920-954-5

Query Match 76.9%; Score 1477.2; DB 3; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1441; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 7 AAGAAAGAGTGTATCTGTATCTGTTTATCAGCTGACGCGATTCTGTCACTGTTCATTA 66
DB 10 AAGAAAGAGTGTATCTGTATCTGTTTATCAGCTGACGCGATTGTGCACTGTTCGTTA 69
QY 67 AACAATCCCTCGCTGTGTGATGCAAGGACTTTTGATCTGGATTTTAAAGAAATTCAGCA 126
DB 70 AGTAATTCATCTGACAGGTGTGTGCAAGAAATTTGATCTGGATTTCAGAAAGAAATTCAGCA 129
QY 127 ACAACGATGTGATGTTTCTCCAAACAGCGAACAAGTGTGCGGCTGATTTCTGTG 186
DB 130 ACAATGATGCTAAAGTTTCTCCAAAGCGGCGAGACTGTGTCTGTCTTTCTGTG 189
QY 187 GAGTCTGAAGATGTGAACCTTCTTAAGGATGCTTAAGAAACCTTGAACAGTACGCGCA 246
DB 190 GAATCTGAAGATGTGAACCTCCCAAGGTTTGCAAGAAAGCTTGAACAGTCCGCGCA 249
QY 247 AATAATAATCTCCATATGTGTCATTCATATGCGCCCATTTTGAAGAAACAAACAGAG 306
DB 250 AATAATAATCTCCATATATTCATTCATATGAGCAAAATTTTGAAGAAACAAACAGAG 309
QY 307 CTAGAGCAACTGAGCAAAAGATCTCGACTACATCCCTGATTATGATATATGTGCGAG 366
DB 310 CTGGAAGAAACAGGCGCAAGAGTTCTCGACTACATCTGATATATGCTTACATTTGTGAG 369
QY 367 TATGAGGGGAGTGTCACTCAAAAGTCCGCTTCATTTGAACACAGTGAATCAGTGAAGCA 426
DB 370 TATGAGGGGAGTGTCACTCAAGCAACAGTTCATTTGAGCAAGTGAATCCGTGAAGCT 429
QY 427 TACTTGCAGAAATCAAAATATGATCCCAAGCTTTTCAAAAAGGCGCATCCAGCTGTG 486
DB 430 TATTTGCGATATACGAATATGATCCCAAGCTTTTCAAAAAGGCGCATCCAGAGCTTTTA 489
QY 487 AAAGGCTTGGCGCTGATACAGAGCAACAATAAAGAGTCAATTAAGAGCATTCGAG 546
DB 490 AAAGAGTGTGGCGCTGATACAGAGCAAAAATTAAGAGTCAATTAAGAGCATTCGAG 549
QY 547 GAATGCTCTAGTACGTAGCAAGCAATGACGTTCATTAATTCAGCAAGGCTGAAATAT 606
DB 550 CAATTCGACATTCGCAATAGCAATGATGTATATTAATTCAGCAAGGCTGAAATAT 609

QY 607 AAGGTATGATGATGTGGCCAGAGGTATTGTCAAAAGCGATGTGGCAAGAGGACCTAC 666
DB 610 AAGGTATGATGATGTGTGGCCAGAGGTATTGTCAAAAGCGATGTGGCCAGAGGACCTAC 669
QY 667 GATTGTATGACAAAGCCAGATGTGTGCACTTCCGATCTGATCTGATGATGATACAGAGA 726
DB 670 GGGTTTATGACAAAGGACAGATGTGTGCGGTTTCCGATACAGAGGCTTGTATACAGGTGC 729
QY 727 AACGACGTTGATGATGAAGCTTCCGCGGTAAATTAACAGCATATATGACTGTGGT 786
DB 730 AATGACGTTGATGATGAAGCTTCCGCGGAAATTAATCTGATTAATATGCAATGGGA 789
QY 787 CGAGCAATTAATGCGAATGATACAGATGATATGATCCCATGAGGAGCTGTGGTATTA 846
DB 790 CGAGCAATTAATGCGAATGATACAGATGATATGATGATGATGATGATGATGATGATGAT 849
QY 847 GGAATGCGCGAAGAAATTAAGAAATGCGCACTCAAGCGAATCTGTTTCAATCCATC 906
DB 850 GGAATGCGCGCAAGAAATTAAGAAATGCGCACTCAAGCGAATCTGTTTCAATCCATC 909
QY 907 ATGATATGAGTGTGTGGCTTGAAGCTTGCCTTCATCTGCAATCTTATTCAGCCAA 966
DB 910 ATGATATGAGTGTGTGGCTTGAAGCTTGCCTTCATCTGCAATCTTATTCAGCCAA 969
QY 967 GCATTCAGTGTGAGTGTGCGAAGTTCATCAAACTCTGGGGGGGAGCGGTGAATGGGGCC 1026
DB 970 GCATTCAGTGTGAGTGTGCGAAGTTCATCAAACTCTGGGGGGGAGCGGTGAATGGGGCC 1029
QY 1027 TACACGACAGATTCACAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
DB 1030 TACACGACAGATTCACAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
QY 1087 TTGCGGCTGTGGAATGAAGGCGCAAGCGGCTACATCACTGATGATCTGTGATGATGAT 1146
DB 1090 TTGCGGCTGTGGAATGAAGGCGCAAGCGGCTACATCACTGATGATCTGTGATGATGAT 1149
QY 1147 AAGCGCAATTAAGTGTGGGCAAGCGAAGCTGTGCTCAAGCTCGATTCCTATGAGAT 1206
DB 1150 AAGCGCAATTAAGTGTGGGCAAGCGAAGCTGTGCTCAAGCTCGATTCCTATGAGAT 1209
QY 1207 AATATTAACAAGTGTGCAAGTCTCTCCGCTGCGCCGCAAAAGATGAGCGATCAAG 1266
DB 1210 AATATTAACAAGTGTGCAAGTCTCTCTTCAAGTGTGCAAGTGTGCAAGTGTGCAAG 1269
QY 1267 CCGATGTCAATGAGCGGCAAGCGGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAG 1326
DB 1270 CCGATGTCAATGAGCGGCAAGCGGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAG 1329
QY 1327 TCCCTCTTCTGGGGGAATGATGACAGCAAAATATGCTATATGAGTGAAGTGCATGGCA 1386
DB 1330 TCCCTCTTCTGGGGGAATGATGACAGCAAAATATGCTATATGAGTGAAGTGCATGGCT 1389
QY 1387 ACACCGAATGTGTGGGGAATGTGCAAGCTCCGTAAGCATTTTGTGAAGAAATGAGAGA 1446
DB 1390 ACACCGAATGTGTGGGGAATGTGCAAGCTCCGTAAGCATTTTGTGAAGAAATGAGAGC 1449
QY 1447 ATCACTCTTAAGCTTCCCTATTAAGAAAGCACTTTGATGAGTGTCTGTGATGATGAT 1506
DB 1450 ATCACTCTTAAGCTTCCCTATTAAGAAAGCACTTTGATGAGTGTCTGTGATGATGAT 1509
QY 1507 TTTGGGTTATCCGGAACGGAACCAAGATGTGGGCGGAGTGAACCCGTGAATATGATGAT 1566
DB 1510 TTTGGGTTATCCGGAACGGAACCAAGATGTGGGCGGAGTGAACCCGTGAATATGATGAT 1569
QY 1567 GTTGCCTATGTGAAGCAATCAAGTGTGCTATCACTAAGCCAAAGGACATATATCTTT 1626
DB 1570 GTTGCCTATGTGAAGCAATCAAGTGTGCTATCACTAAGCCAAAGGACATATATCTTT 1629
QY 1627 ACTGCAACGCGGCGCAAGCATTTGAATATCTCTGTGATGTGTGATGATGATGATGATGAT 1686
DB 1630 ACTGCAACGCGGCGCAAGCATTTGAATATCTCTGTGATGTGTGATGATGATGATGATGAT 1689

```
OY 1687 ACTACGCTCTGTACCCGTGTCATGATTTGATTTGTGATTCATACGACCAAAACGGA 1746
|||
DB 1690 ACACCTGCTCCGTAAACGCTGTCAATGATCTGACCTTGATACCGCTCAAAATGAC 1749
OY 1747 ACAGATGTCGGGAAATGATCTTCTCAGACCAATTTGACATTAATGAGATGGCGCAT 1806
|||
DB 1750 ACACGATGATGAGAAATGATCTTCTGCTGCTACATGATGATGATGAGATGGCGCAT 1809
OY 1807 AACGTGAAATGATATTTATTTATTTGCGCCCAAGTGAAATATACATTTAGATGTCGA 1866
|||
DB 1810 AACGTGAAATGATATTTATTTATTTATGACCAAAAGGGGACGTATACATTTAGATGTCAG 1869
OY 1867 GCATATATGTCGCGTGGACCAAACTTCTGTTGGCAATTTGTAATCTTA 1920
|||
DB 1870 GCTTAAACGTACCGGTGACCAACAGACTTCTGTTGGCAATTTGTAATTTAA 1923

RESULT 4
US-09-920-954-7
; Sequence 7, Application US/09920954
; Publication No. US20020064854A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/J998/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-09-920-954-7

Query Match 76.9%; Score 1477.2; DB 3; Length 1923;
Best Local Smilarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

OY 7 AAGAAAGAGGTGTTTATCTGTTTATCAGTCGACGAGATTCGTGCACTGTTGCATTA 66
|||
DB 10 AAGAAAAGGTGTTTATCTGTTTATCAGTCGACGAGATTTGTCACATGTTGCGTTA 69
OY 67 AACATCCCTGGGCTGATGATGCAAGGACTTTGATCTGATCTGATTTTAAAGGAATTAACA 126
|||
DB 70 AGTATTCATCTGACGATGATGCAAGGATTTTATGATCTGATTTTAAAGGAATTAACA 129
OY 127 ACAACGATGTCAGTGTGTTTCTCAAAAGCAAGAGATGCGGCTGCAATTTCTGATG 186
|||
DB 130 ACACTGATGCTAAAGTTTCTTCCAAAGCAGGCGCAGATGCTGCTGCTCTTCTGATG 189
OY 187 GAGTGTAAATGTAATCTTCTTAAAGATTTGCTAAAGAACTTGAACAGTACCGGCA 246
|||
DB 190 GAATCTGAAAATGTAATCTCCCAAAAGGTTTGCAAGAAAGCTTGAACAGTCCGGCA 249
OY 247 AATATATACTCCATATTTGCAATTTCAATGCGCCCAATTTTAAAGAAACAAACAGAG 306
|||
```

```
DB 250 AATATATACTCCATATTTATCCATTCATGACCAATTTTGAAGAAACAAACAGCAG 309
OY 307 CTAGAGCAACTGAGCAAAAGATTTCTGACTACATCTCCGATTAATGATATTTGTGAG 366
|||
DB 310 CTGAAAAAAGAGGAGCAAAAGATTTCTGACTACATACCTGATTAATGATTTGTGAG 369
OY 367 TATGAGGGGATGTTCACTCAAAAGTCCGCTCATTTGAACAGCTGATTCAGTGAAGCA 426
|||
DB 370 TATGAGGGGATGTTGAATCAGCAAAAGCAACATTTGAGCAGTGAATCCGTGAGCCT 429
OY 427 TACTGCGGAAATACAAATATGATCCGACTTTTCAAAAGGGGCAATGACGCTGATG 486
|||
DB 430 TATTTGCGGATATACAAATATGATCCGACTTTTCAAAAGGGGCAATGAGCTGTA 489
OY 487 AAAGGTTGCGCTTGAATACGAAGCAACATTAAGAATGTCAAATTAAGAGCATGAG 546
|||
DB 490 AAAGAGTGGCGCTTGAATACGAAGCAAAATTAAGAGTGCATTAAGAGCATGAG 549
OY 547 GAAATCGCTCAGTACGTAGCAAGCAATGACCTCATTTATTTAGCGCAAGCTGATAT 606
|||
DB 550 CAATATGCACAATTTGCAATTAAGCAATGATGCTATATTAATTAACGCAAGCTGATAT 609
OY 607 AAGGATGAATGATGTCGCGGCAAGGATATTTGCAAGGGAGTGGCAAGAGCATAC 666
|||
DB 610 AAGGTGAATGATGTTGCGCGGAAATTTGCAAGGGAGTGGCTGACAGCATAC 669
OY 667 GGTGTGATGCAAGGCGCAGATTTGTCAGTTGCCGATCTGATGATGATGATCAGAGAA 726
|||
DB 670 GGGTTGATGCAAGGCAAGCAAGTCTGACGTTGCGATACAGGGCTTGAATCAGGTGCG 729
OY 727 AACGACGTTGATGATGAAGCTTCCGCGGTAATTAACGACATATATGACCTGGT 786
|||
DB 730 AATGACGTTGATGATGAAGCTTCCGCGGTAATTAACGACATATATGACCTGGT 789
OY 787 CGGAGGAATTAATGGAATGATGACGATGATGATGATGATGATGATGATGATGATGAT 846
|||
DB 790 CGGAGGAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
OY 847 GGAATGCGGCAACCAATTAAGAAATGCACTCAAGGCAATCTGTTTTCATTCATC 906
|||
DB 850 GGAATGCGGCAACCAATTAAGAAATGCACTCAAGGCAATCTGTTTTCATTCATC 909
OY 907 ATGATGACAGTGTGCGCTTGAAGGCTTGCCTTCAATCTGCAAACTTATTCAGCCA 966
|||
DB 910 ATGATGACAGTGTGCGCTTGAAGGCTTGCCTTCAATCTGCAAACTTATTCAGCCA 969
OY 967 GCATTCAGTGTGCGCTTGAAGGCTTGCCTTCAATCTGCAAACTTATTCAGCCA 1026
|||
DB 970 GCATTCAGTGTGCGCTTGAAGGCTTGCCTTCAATCTGCAAACTTATTCAGCCA 1029
OY 1027 TACACGACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
|||
DB 1030 TACACGACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
OY 1087 TTCCGCGCTGGAATGAAGGCGCAACGCGGATCAATCAGTGCACCTGTAACGCTTAA 1146
|||
DB 1090 TTCCGCTGCGGGAATGAAGGCGCAACGCGGATCAATCAGTGCACCTGTAACGCTTAA 1149
OY 1147 AACGCAATTAAGTGTGCGGCAACGGAACCTGCGTCCAACTTCGTTCTTATGCAAT 1206
|||
DB 1150 AATGCAATTAAGTGTGCGGCAACGGAACCTGCGTCCAACTTCGTTCTTATGCGGAC 1209
OY 1207 AATATTAACAAGTTGCAAGTTCCTTCCGCTGCGGCGCAAAAGATGAGGCGAATCAAG 1266
|||
DB 1210 AATATTAACAAGTTGCAAGTTCCTTCCGCTGCGGCGCAAAAGATGAGGCGAATCAAG 1269
OY 1267 CCTGATGTCATGCGCGCAAGCAATCAATTTATCAAGCAATCTTCTTTCAGCCGAT 1326
|||
DB 1270 CCTGATGTCATGCGCGCAAGCAATCAATTTATCAAGCAATCTTCTTTCAGCCGAT 1329
OY 1327 TCTCTCTTCTGCGGGAATCAAGCAAGCAAAATGCTTATGAGGTTGGAACCTTCATGCA 1386
|||
DB 1330 TCTCTCTTCTGCGGGAATCAAGCAAGCAAAATGCTTATGAGGTTGGAACCTTCATGCA 1389
|||
```

Qy 1387 AACGATTGTTGGGGGAATGTCAGAGCTCCGTGAGCAATTTTGTGAAAAATAGAGA 1446
Db 1390 AACGATTGTTGGGGGAATGTCAGAGCTCCGTGAGCAATTTTGTGAAAAATAGAGA 1449
Qy 1447 ATCACTCTTAAGCCCTTCCCTATGTAAGCAGCTTGTGATGAGTGTGCTGATGTTGA 1506
Db 1450 ATCAACACCAAGCCCTTCTATTAAGAGGAGCACTGATTTGGCGGTGAGCTGACATCGGC 1509
Qy 1507 TTGGGTTATCCGAAAGCAAGCAAGAGTGGGCGAGTGAACCTTGATTAATCGTTGAAC 1566
Db 1510 CTGGCTACCCGAAAGCAAGCAAGAGTGGGCGAGTGAACCTTGATTAATCGTTGAAC 1569
Qy 1567 GTTGCCTATGTAAGCAATCCAGTGCCTATCACTAGCAAAAGCAATATACCTTT 1626
Db 1570 GTTGCCTATGTAAGCAATCCAGTGCCTATCACTAGCAAAAGCAATATACCTTT 1629
Qy 1627 ACTGCAACGCGCGGCAAGCAATGTAATACTCCCTGTATGATGCGATGCCCTGCAAGC 1686
Db 1630 ACTGCTATCGCGGCAAGCAATGTAATACTCCCTGTATGATGCGATGCCCTGCAAGC 1689
Qy 1687 ACTACTGCTTCTGTAACTCTGTGATGATGTTGATTTGATTCAGACCAAGCA 1746
Db 1690 ACAACTGCTTCTGTAACTCTGTGATGATGTTGATTTGATTCAGACCAAGCA 1749
Qy 1747 ACAAGATATGTCGGGATGACTTCTCAGACCACTTGAACAATACCTGGAGTGGCGCAAT 1806
Db 1750 ACAAGATATGTCGGGATGACTTCTCAGACCACTTGAACAATACCTGGAGTGGCGCAAT 1809
Qy 1807 AACGTAGAAATGTAATTTATTAATTCGCCCCCAAGTGAACATATACCTGAGTGCNA 1866
Db 1810 AACGTAGAAATGTAATTTATTAATTCGCCCCCAAGTGAACATATACCTGAGTGCNA 1869
Qy 1867 GCATATATGTCGGGATGACTTCTCAGACCACTTGAACAATACCTGGAGTGGCGCAAT 1920
Db 1870 GCATATATGTCGGGATGACTTCTCAGACCACTTGAACAATACCTGGAGTGGCGCAAT 1923

RESULT 5

US-10-456-479-3
; Sequence 3, Application US/10456479
; Publication No. US20040072321A1
; GENERAL INFORMATION:
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: TAKIMURA, YASUSHI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: NOMURA, MASAFUMI
; APPLICANT: KOBAYASHI, TOHRU
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 238700US0
; CURRENT APPLICATION NUMBER: US/10/456,479
; PRIOR FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: JP 2002-186387
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: JP 2002-304232
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 3
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-KP43
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1920)
; OTHER INFORMATION:
US-10-456-479-3

Query Match 76.9%; Score 1477.2; DB 7; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Qy 7 AAGAGAGGTGTTTTTATCTGTTTTATCAGTCAAGCAATTTCTGTGCACTGTGCATTA 66
Db 10 AAGAGAGGTGTTTTTATCTGTTTTATCAGTCAAGCAATTTCTGTGCACTGTGCATTA 69
Qy 67 AACATCCCTGGCTGTGTGATGCAAGCACTTTGATCTGATATTTTAAAGAAATTCAAACA 126
Db 70 AGTATCCATCTGCAAGGTGTGCAAGCAATTTTATGATCTGATATTTCAAGAAATTCAGACA 129
Qy 127 AACACCAATGTCAGTGTGTTTTCTCAACACAGCAACAAACAGTGGCGTGCATTTCTGTG 186
Db 130 AACACCAATGTCAGTGTGTTTTCTCAACACAGCAACAAACAGTGGCGTGCATTTCTGTG 189
Qy 187 GAGTCTGAAATGTAACCTTTAAAGAAATGCTTAAAGAAATCAGTACCGCA 246
Db 190 GAGTCTGAAATGTAACCTTTAAAGAAATGCTTAAAGAAATCAGTACCGCA 249
Qy 247 AATATTAATCTCAATATGTCCAATTCATATGAGCCCATTTTAAAGAAACAAACAGAG 306
Db 250 AATATTAATCTCAATATGTCCAATTCATATGAGCCCATTTTAAAGAAACAAACAGAG 309
Qy 307 CTAGAGCACTGAGCAAGCAATTTCTGACATCAATCCCTGATTAATGATATGTCAG 366
Db 310 CTAGAGCACTGAGCAAGCAATTTCTGACATCAATCCCTGATTAATGATATGTCAG 369
Qy 367 TATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCAATGCAACGTGCAATCAGTGAAGCA 426
Db 370 TATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCAATGCAACGTGCAATCAGTGAAGCA 429
Qy 427 TACTTCCGAAATACAAATATGATCCCAAGCTTTTCAAAAGGCGCATGCACTGTGTG 486
Db 430 TACTTCCGAAATACAAATATGATCCCAAGCTTTTCAAAAGGCGCATGCACTGTGTG 489
Qy 487 AAGCGTTGGCGCTGATTAAGCAAGCAATTAAGAGTGAATTAAGGGGATGAG 546
Db 490 AAGCGTTGGCGCTGATTAAGCAAGCAATTAAGAGTGAATTAAGGGGATGAG 549
Qy 547 GAAATCCCTCAGTACGTAGCAAGCAATGACCTTCATTAATTAAGCAAGGCTGAAATAT 606
Db 550 GAAATCCCTCAGTACGTAGCAAGCAATGACCTTCATTAATTAAGCAAGGCTGAAATAT 609
Qy 607 AAGGTATGAATGATGTCGCAAGGATGATTTCTCAAAGCGGATGTCACAGAGCACTAC 666
Db 610 AAGGTATGAATGATGTCGCAAGGATGATTTCTCAAAGCGGATGTCACAGAGCACTAC 669
Qy 667 GCTTGTATGCAAGCAAGCAATGTCGCAAGGATGATTTCTCAAAGCGGATGTCACAGAGCACTAC 726
Db 670 GCTTGTATGCAAGCAAGCAATGTCGCAAGGATGATTTCTCAAAGCGGATGTCACAGAGCACTAC 729
Qy 727 AACGACAGTTCGATGCAATGAGCACTTCGCGGTAAATTAACAGCACTATATGCACTGAGT 786
Db 730 AACGACAGTTCGATGCAATGAGCACTTCGCGGTAAATTAACAGCACTATATGCACTGAGT 789
Qy 787 CGAGCAATTAATGCAATGATACGAAGTCAATGATACCAATGTCGCAAGTTCGATATTA 846
Db 790 CGAGCAATTAATGCAATGATACGAAGTCAATGATACCAATGTCGCAAGTTCGATATTA 849
Qy 847 GGAATATGCGCAAGCAATTAAGCAAGTCAATGATACCAATGTCGCAAGTTCGATATTA 906
Db 850 GGAATATGCGCAAGCAATTAAGCAAGTCAATGATACCAATGTCGCAAGTTCGATATTA 909
Qy 907 ATGATAGCAGTGTGAGCTTGAAGCTTGCCTTCAATGCAATGCAATCTTATGAGCA 966
Db 910 ATGATAGCAGTGTGAGCTTGAAGCTTGCCTTCAATGCAATGCAATCTTATGAGCA 969
Qy 967 GCATTCAGTGCAGTGTGCAAGATTCATACAACTCTGAGGGGCAAGGATGAGTGGGCT 1026
Db 970 GCATTCAGTGTGAGCTTGAAGCTTGCCTTCAATGCAATGCAATCTTATGAGCA 969
Qy 1027 TACACGACATTCACAAATATGTCAGTATGATGATGATGATGATGATGATGATGATGAT 1086
Db 1030 TACACGACATTCACAAATATGTCAGTATGATGATGATGATGATGATGATGATGATGAT 1089
Qy 1087 TTCGCGCTGGGAATGAAGCGCAAGCGGCTACATCACTGACCTGTGATCGGCTTAA 1146

OY	787	CGGAGGAATAATGCGGAATGATATCGAAACGGTCATGATGCCAATGCGCAGGTTGCGATTTA	846
Db	790	CGGAGGAATAATGCGCATATGATGAAATGGTCATGATCGCATGATGGCTCCGATTTA	849
OY	847	GGAAATGCGCGCAACGAATTAAGGAATGGCACTCAAGCGAATCGTGTTTTTCATCATC	906
Db	850	GGAAACGGCTCCACTATATTAAGGAATGGCGCCTCAAGCGAATCTAGCTTCCATCTATC	909
OY	907	ATGATAGCAGTGTGTGGGCTTGGAGGCTTGCTTCCATCTGCAACCTTATTCAGCGAA	966
Db	910	ATGATAGCAGGTGGGGGAGCTTGGAGGAGCTACTTCGAATCTGCAAACTTATTCAGCGAA	969
OY	967	GGAATTCAAGTCAGTGTGCCGAATTCATATACAAATCTCTGGGGGGGCAACGGTGAATGGGGCC	102
Db	970	GGAATCAAGTCAGTGTGCCGAATTCATATACAAATCTCTGGGGGGGCAACGAATGGGGCT	102
OY	1027	TACACGACAGATTTCCAGAAATGTGATGACTATGTAAGGAATAATGATATGACGATTTCTT	108
Db	1030	TAACACGACAGATTTCCAGAAATGTGATGACTATGTCGCAAAAATGATATGACGATTTCTT	108
OY	1087	TTGCGGGCTGGGAATGAAGCCGGAACGGCGGTACATCAGTGCACCTGTGTACGGCTTAA	114
Db	1090	TTGCGTGCAGGGAATGAAGCCGGAACCGGAGCCATCAGTGCACAGGCTTAA	114
OY	1147	AAACGCTATACAGTGTGGCGCCAAACCGAATACTGTGCTCAAGCTTCGTTCTTATGCAAT	120
Db	1150	AAATGAATTAACAGTGTGGAGCTACGGAATACTTCGCGCAAGCTTGTGTATATGGCGAAC	120
OY	1207	AATATTAAACCAAGTTTGCAAGTTCCTTTCGCGGGCCGCACAAAAGATGGCGAATCAAG	126
Db	1210	AATATCAACCAATGTGGCAAGTTCTTTCACTGTGACCGACAAAGATGGACGGAATCAAA	126
OY	1267	CTGTATGTATGAGCGCGCAAGGACATCATTTTATCAGCAAGATCTTCTTTCGACCCGAT	132
Db	1270	CCGAGATGTATGAGCAACCGGGAACGTTCAATCATCAGCAAGATCTTCTTTCGACCGGAT	132
OY	1327	TCCTCTCTCTGTGGCGCAATCATGACAGCAATAATGCTTATATGGTGTGAACGTCTCATGGCA	138
Db	1330	TCCTCTCTCTGTGGCGCAACCATGACAGTAAATATGACATGGTGTGAACGTCTCATGGCT	138
OY	1387	ACACCGAATTGTGCGGGGAATGTTGACAGCTCCGAGAGCATTTTGTGAAATAATAGAGA	144
Db	1390	ACACCGAATGTGTGTGTGGAACGTGGACAGCTTCTGTAGAGCATTTTGTGAAATAACAGAGC	144
OY	1447	ATCAGCTCTTAAGCCTTCCCTATTTGAAGAGCTTTGATTTGCAAGTGTGCTGTATGTTTGA	150
Db	1450	ATCAGCACCMAAGCCTTCTCTATTTAAAGGGGAGCTGATTTGGCGGTGCAAGTGCATGTGG	150
OY	1507	TTGGGTTATCCGAAGGAAACCAAGATGGGGCCGAGTGCCTGTGATTAATCTGTTAAC	156
Db	1510	CTTGGCTATCCCGAACGGTAAACCAAGATGGGAGCGATGACATTTGATTAATCTCCGTAAC	156
OY	1567	GTTGCTTATGTAAGCAATCCAGTGGCCATATCAATAGGCAAAAAGGACATTAACCTTT	162
Db	1570	GTTGCTTATGTAAGCAATCCAGTGGCTTCTATTCACAGCGCAAAAAGCAAGTATCTGTTT	162
OY	1627	ACTGCTAAGCGCGGAGCAAGCATTTGAAATCTCCCTGTATGTCGATGATGCCCTGTGAGC	168
Db	1630	ACTGCTATGCGGGAGAGCCTTTGAAATATCTCCCTGTATGTCGATGATGCCCTGTGAGC	168
OY	1687	ACTATGCTCTTCTGTAAACCTGTGATATTTGGATTGGTCAATTACAGCACCAACGGA	174
Db	1690	ACAATGTGCTTCGTAACGCTTGTCAATGATCTGTGACCTTGTCAATTACCCCTCCAAATGAGC	174
OY	1747	ACAAGATATGTGGGAATGACTTTCAGACCAATTTGAACAATPACGTGGGATGGCCGCAAT	180
Db	1750	ACAAGATATGTGGGAATGACTTTCAGACCAATTTGAACAATPACGTGGGATGGCCGCAAT	180
OY	1807	AAAGTAGAAATGTATTTTATTTATTTTCGCCCCCAAGTGTGAACATATCAATTGAGGTGCA	186
Db	1810	AAAGTAGAAATGTATTTTATTTATTTATGACCAACAAGGGGAGCGTATCAATTGAGGTGCA	186
OY	1867	GCAATTAATGTGCGGTGGAGCACAAACCTTCTGTGGCAATTGTGAATTA	1920

```

Db      1870  GCTTATAACGTAACGGGTGGACCAAGACCTTCTGCTGGCAATTGTGATTTAA 1923
      |||||
RESULT 7
US-10-784-870-7
/ Sequence 7, Application US/10784870
/ Publication No. US20040142837A1
/ GENERAL INFORMATION:
/ APPLICANT: TAKAIWA, MIKIO
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SAEKI, KATSUHIISA
/ APPLICANT: KUBOTA, HIROMI
/ APPLICANT: HITOMI, JUN
/ APPLICANT: KAGEYAMA, YASUSHI
/ APPLICANT: SHIKATA, SHITSUM
/ APPLICANT: NOMURA, MASAFUMI
/ TITLE OF INVENTION: ALKALINE PROTEASE
/ FILE REFERENCE: 0327-0832-0PCT
/ CURRENT APPLICATION NUMBER: US/10/784, 870
/ CURRENT FILING DATE: 2004-02-24
/ PRIOR APPLICATION NUMBER: US/09/509, 814A
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: PCT/JP98/04528
/ PRIOR FILING DATE: 1998-10-07
/ PRIOR APPLICATION NUMBER: JP 9-274570
/ PRIOR FILING DATE: 1997-06-08
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 7
/ LENGTH: 1923
/ TYPE: DNA
/ ORGANISM: Bacillus sp.
/ FEATURES:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1923)
US-10-784-870-7

Query Match      76.9%; Score 1477.2; DB 7; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY      7  AAGAAAGAGGTGTTTTATCTGTTTTATCAGCTGACGAGGATTCGTGCACTGTTCATTAA 66
DB      10  AAGAAAAAGGGTTTTATCTGTTTTATCAGCTGACGAGATTTGTGCACTGTTCGTTA 69
QY      67  AACAAATCCCTCGGTGGATGCAAGGACTTTGATCTGATTTTAAAGAAATTCAAA 126
DB      70  AGTATATCCATTCGACGGGTGGCAAGGAATTTGATCTGATTTTCAAGGAATTCAGCA 129
QY      127  ACAACCGATGTCAGTGGTTTCTCCAAACAGCGCAAAACAGGTGCGGTGATTCGTG 186
DB      130  ACAATGATGCTAAAGGTTTCTCCAAACAGGGGCGAGACTGCTGCTCTTTCTG 189
QY      187  GAGTCTGAAAATGTGAAACTTTCTTAAAGATTTGCTTAAAGAAACTTGAACCGGCA 246
DB      190  GAATCTGAAAATGTGAAACTCCCAAAAGGTTTGCAAGAAAGCTTGAACAGTCCGGCA 249
QY      247  AATATAAATCCCATTTATTTGCCAATTCGATGAGCCCAATTTTAAAGAAAGAAACAGAG 306
DB      250  AATATAAATCCCATTAATATTCGAATTCGATGAGCAATTTTAAAGAAAGAAACAGCAG 309
QY      307  CTAGAGCAACCTGAGCAAAAGATTCTGCACTACATCCCTGATTATGATATAATTGCGAG 366
DB      310  CTGAGAAAAACAGGGGCGAAAGATTCTGCACTACATCTGATTATGCTTACATTGTGAG 369
QY      367  TATGAGGGGAGATTTCACTCAAAAGTCGCTCCATTGAACACGTGGAATCAGTGAAGCCA 426
DB      370  TATGAGGGGAGATTTAAGTCAGCAAAAGCAACCATTTAGACGCGGAATCGTGAGAGCT 429
QY      427  TACTTGGCGAATTCAAAATGAGTCCCGAGCTTTTCAAAAAGGCGCATGACGCTG 486
DB      430  TATTTCGCAATATACAGATGAGTCCCGAGCTTTTCAAAAAGGCGCATGACGCTG 489

```


130 ACAACTGATGCTAAAGTTCCTCCAGAGAGGGGCGAGACTGCTGCTCTTTCTGGTG 189
187 GAGTGTGAAAATGTGAACTTCTTAAAGATGCTTAAAGAACTTGAACAGTACCGGCA 246
190 GAATCTGAAAATGTGAACTCCCAAAAGGTTGCGAAGAACTTGAACAGTCCCGGCA 249
247 AATAATTAACCTCCATATGTCCAAATTCAAATGGCCCCATTTTGAAGAAACAAACAGAG 306
250 AATAATTAACCTCCATATATCCAAATTCAAATGGAACCAATTTTGAAGAAACAAACAGAG 309
307 CTAGAGCAACCTGAGCAAAAGATTCCTGACTACATCCCTGATATATGATATGCTGAG 366
310 CTGAAAAAACAAGGGGCAAAAGTTCTCGACTACATACCTGATATGCTTCACTTGTGAG 369
367 TATGAGGGGAGTGTTCAGTCAAAAGTCCGCTTCATTTGAACAGTGAATCAGTGAAGCCA 426
370 TATGAGGGGAGTGTTCAGTCAAAAGTCCGCTTCATTTGAACAGTGAATCAGTGAAGCCA 429
427 TACTTGCAGAAATCAAAATGATGCTCCGAGCTTTTCAAAAAGGGGCAATCCAGCTGTG 486
430 TATTTGCCGATATACAAATGATGATCCGAGCTTTTCAAAAAGGGGCAATCCAGCTGTG 489
487 AAAGCTTGGCGCTGATACGAGAGCAACATTAAGAAGTCAATTAAGAAGCATCGAG 546
490 AAAGCAGTGGCGCTGATACGAGAGCAACATTAAGAAGTCAATTAAGAAGCATCGAG 549
547 GAATCGCTCAGTACGTAAGCAAGCAATGACCTTATATATTAATTAATTAATTAATTAAT 606
550 CAATTCGCAATTCGCAATTAAGCAATGATGCTATATATTAATTAATTAATTAATTAAT 609
607 AAAGGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
610 AAAGGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
667 GATTGTATGAGCAAGCGCAGATGTGCGCAGTTGCGGATACCTGATGATGATGATGATGAT 726
670 GGGTTGTATGAGCAAGCGCAGATGTGCGCAGTTGCGGATACCTGATGATGATGATGATGAT 729
727 AACGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
730 AATGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
787 CGGAGCAATTAATGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
790 CGGAGCAATTAATGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
847 GGAATGCGCGCAACGATTAAGAAATGCGACCTCAAGCGAATCTGTTTCAATCCATC 906
850 GGAATGCGCGCTCACTAATTAAGAAATGCGCGCTCAAGCGAATCTGTTTCAATCCATC 909
907 ATGATGACAGTGTGCTGAGGCTTGGAGGCTTGGCTTCAATCTGCAAACTTATTCAGCCAA 966
910 ATGATGACAGTGTGCTGAGGCTTGGAGGCTTGGCTTCAATCTGCAAACTTATTCAGCCAA 969
967 GCATTCAGTGAAGGCGAGAAATCACTCAAACTCTGAGGGGGGCGAGGCTGAATGGGGCC 1026
970 GCATTCAGTGTGCTGAGGCTTGGAGGCTTGGCTTCAATCTGCAAACTTATTCAGCCAA 1029
1027 TACACGACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
1030 TACACGACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
1087 TTGCGGCTGTGAGAAATGAAGCGGAGCGGCTGATCAGTGAAGCTGCTGCTGCTGCTGCT 1146
1090 TTGCGTGTGCGGAAATGAAGCGGAGCGGAGCAGTGAAGCTGCTGCTGCTGCTGCTGCT 1149
1147 AACGCGATTAAGTGTGCGGAGCAACGAAACCTGCGTCAAGCTTGGCTGCTGCTGCTGCTGCT 1206
1150 AATGCGATTAAGTGTGCGGAGCAACGAAACCTGCGGCGCAAGCTTGGCTGCTGCTGCTGCTGCT 1209
1207 AATATTAACAGATTTGCAAGTTCCTTCCGCTGAGCCGCAAAAGATGGGCGATCAAG 1266
1210 AATATTAACAGATTTGCAAGTTCCTTCCGCTGAGCCGCAAAAGATGGGCGATCAAA 1269

1267 CCGATGTATGAGGCGCCAGGAGCATACATTTTATCAGCAAGATCTTCTTGGACCCGAT 1326
1270 CCGATGTATGAGGCGCCAGGAGCATACATTTTATCAGCAAGATCTTCTTGGACCCGAT 1329
1327 TCTCTCTTCTGGGGAATCATGACGCAATATGCTATATGAGTGAACGCTCATGAGCA 1386
1330 TCTCTCTTCTGGGGAATCATGACGCAATATGCTATATGAGTGAACGCTCATGAGCT 1389
1387 ACACCGATGTTGGCGGGAATGTTGCAAGCTCCGTAAGCATTTTGTGAAAAATGAGGA 1446
1390 ACACCGATGTTGGCGGGAATGTTGCAAGCTCCGTAAGCATTTTGTGAAAAATGAGAGC 1449
1447 ATCACTCTTAAGCTTCTCTATTTGAAAGCAGCTTGTGATGAGTGTGCTGATGTTGGA 1506
1450 ATCACTCTTAAGCTTCTCTATTTGAAAGCAGCTTGTGATGAGTGTGCTGATGTTGGA 1509
1507 TTGGGTTATCCGAAACGGAACCAAGATGAGGCGGAGTGAACCTGATTAATTCGTTGAGC 1566
1510 CTGAGCTACCCGAAACGTTAACCAGATGAGGAGGAGTGAACATTTGATTAATTCCTGAGC 1569
1567 GTTGCTATGTAAGCAATTCAGTGCCTATCACTAGCCAAAGCGACATATACCTTT 1626
1570 GTTGCTATGTAAGCAATTCAGTGCCTATCACTAGCCAAAGCGACATATACCTTT 1629
1627 ACTGCAACGCGGCGAGCCATTTGAAATCTCCCTGATGATGATGATGATGATGATGATGAT 1686
1630 ACTGCTACTGCGGCGAGCCATTTGAAATCTCCCTGATGATGATGATGATGATGATGATGAT 1689
1687 ACTACTGCTCTGTAACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1746
1690 ACAACTGCTCTGTAACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1749
1747 ACAAGATATGTCGGAATGACTTCTCAGACCAATTTGACATTAATGAGGATGCGCAT 1806
1750 ACAAGATATGTAAGAAATGACTTACTTCCGATACATATGATTAATGAGGATGCGCAT 1809
1807 AACGTAAAAATGATTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1866
1810 AACGTAAAAATGATTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1869
1867 GCATATATGTCGCGGTTGAGCAACAAACTTCTGTTGCAATTTGATTA 1920
1870 GCTTATTAACGTAACCGGTTGAGCAACAACTTCTGTTGCAATTTGATTA 1923

RESULT 9
US-10-820-714A-2
; Sequence 2, Application US/10820714A
; Publication No. US20050214922A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuyoshi
; APPLICANT: Kobayashi, Toku
; APPLICANT: Sumitomo, Nobuyuki
; APPLICANT: Takimura, Yasushi
; APPLICANT: Sato, Tsuyoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 251697US0
; CURRENT APPLICATION NUMBER: US/10/820,714A
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106709
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-KP43
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1920)
; FEATURE:

NAME/KEY: s19_peptide
LOCATION: (1)...(618)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (619)...
US-10-820-714A-2

Query Match 76.9%; Score 1477.2; DB 9; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 7 AAGAAAGAGTGTGTTTATCTGTTTATCAGTCGACGCAATTCGTGACTGTGATTA 66
DB 10 AAGAAAGAGTGTGTTTATCTGTTTATCAGTCGACGCAATTCGTGACTGTGATTA 69
QY 67 AACATTCCTCGGCTGCTGATGACAGCACTTTGATCTGATTTTAAAGAAATTCAGCA 126
DB 70 AGTAATCCATCTGACGCTGCTGACAGGAATTTGATCTGATTTCAAAGAAATTCAGCA 129
QY 127 ACAACGATGTGAGTGTGTTTCTCCAAACAGCGCAAAACGCTGCTGCAATTTCTGTG 186
DB 130 ACAACGATGTGAGTGTGTTTCTCCAAACAGCGCGAGAGCTGTGCTGCTGCTTTCTGTG 189
QY 187 GAGTCGAAATATGTGAACTTCTTAAAGGATGCTTAAAGAACTTGAACAGTACCGGCA 246
DB 190 GATCTGAAATATGTGAACTTCTCCAAAGGTTTGCAGAGAGAGCTTGAACAGTCCGCGCA 249
QY 247 AATAATTAATCTCATATTTGTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 306
DB 250 AATAATTAATCTCATATTTGTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 309
QY 307 CTAGAGACAACTGAGGCAAAAGATTTCTGCACTACATCCGATTTATGATTTATGTCGAG 366
DB 310 CTGAGAAACAGGCGCAAGATTTCTGCACTACATTCATTCATTCATTCATTCATTCATTC 369
QY 367 TATGAGGAGGATGTTCACTCAAAAGTCCGCTCATTTGAACAGTGGATCAGTGAAGCA 426
DB 370 TATGAGGAGGATGTTAAGTCAAGCAACAGACCACTTGAACAGTGGATCAGTGAAGCT 429
QY 427 TACTGCGGAAATCAAAATAGATCCCAAGCTTTTCAAAAAGCGCATGACGCTGTG 486
DB 430 TATTTGCCGATTAAGAAATAGATCCCAAGCTTTTCAAAAAGCGCATGACGCTGTG 489
QY 487 AAAGGCTGCGCTGATTAAGAGCAACAAATTAAGAGTCAATTAAGAGCATTCGAG 546
DB 490 AAAGGCTGCGCTGATTAAGAGCAACAAATTAAGAGTCAATTAAGAGCATTCGAG 549
QY 547 GAAATCGCTCAGTACGTAAGCAATGACGTCATTTATTAACGCAAAAGCTGATAT 606
DB 550 GAAATCGCTCAGTACGTAAGCAATGACGTCATTTATTAACGCAAAAGCTGATAT 609
QY 607 AAGGTATGATGATGTGCGCAAGAGTATGTCAAAGCGAGTGTGCAACAGCACTAC 666
DB 610 AAGGTATGATGATGTGCGCGTGAATTTGCAAAAGCGAGTGTGCGTCAAGCACTAC 669
QY 667 GGTGTTGATGACAGGCGCAATTTGTCGATGCTGCGATCTGGAATGGATACAGGAAGA 726
DB 670 GGTGTTGATGACAGGCGCAATTTGTCGATGCTGCGATCTGGAATGGATACAGGTCGC 729
QY 727 AACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
DB 730 AATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 789
QY 787 CGAGCAATTAATGCAATGATACGAGCTGATGATGATGATGATGATGATGATGATGATG 846
DB 790 CGAGCAATTAATGCAATGATACGAGCTGATGATGATGATGATGATGATGATGATGATG 849
QY 847 GGAATGCGCGCAAGCAATTAAGGATGCGCACTCAAGCAATCTGTTTTCATTCATC 906
DB 850 GGAATGCGCGCTCAATTAAGGATGCGCTCAAGCAATCTGTTTTCATTCATTCATC 909
QY 907 ATGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 966

DB 910 ATGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 969
QY 967 GCAATCACTGCAAGTGCAGCAATTCATTAACATCTCTGCGGCGCAGCGGTGAATGCGGCC 1026
DB 970 GCAATCACTGCGTGCAGCAATTCATTAACATCTCTGCGGCGCAGCGGTGAATGCGGCC 1029
QY 1027 TACAGCAAGATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCA 1086
DB 1030 TACAGCAAGATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCA 1089
QY 1087 TTGCGGCTGCGCAATTAAGGCGCAAGCGCGTACCATCATGATGATGATGATGATGATGAT 1146
DB 1090 TTGCGGCTGCGCAATTAAGGCGCAAGCGCGTACCATCATGATGATGATGATGATGATGAT 1149
QY 1147 AACGCAATTAAGTGCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCA 1206
DB 1150 AATGCAATTAAGTGCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCA 1209
QY 1207 AATATTAACAGTGTGCAAGTTCCTTCCGCTGCGCGCAAAAGATGATGATGATGATGAT 1266
DB 1210 AATATTAACAGTGTGCAAGTTCCTTCCGCTGCGCGCAAAAGATGATGATGATGATGAT 1269
QY 1267 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
DB 1270 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1329
QY 1327 TCTTCTTCTGCGGCAATCATGACAGCAATTAAGCTTAATGATGATGATGATGATGATGAT 1386
DB 1330 TCTTCTTCTGCGGCAATCATGACAGCAATTAAGCTTAATGATGATGATGATGATGATGAT 1389
QY 1387 ACAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
DB 1390 ACAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1449
QY 1447 ATCACTCTTAAGCTTCTCTTATTAAGCAAGCTTATTAAGCAAGCTTATTAAGCAAGCTT 1506
DB 1450 ATCACTCTTAAGCTTCTCTTATTAAGCAAGCTTATTAAGCAAGCTTATTAAGCAAGCTT 1509
QY 1507 TTGCGTATTCGCAACGAAACCAAGATGCGCGCGAGTGAACCTTGATTAATGCTTGAAC 1566
DB 1510 TTGCGTATTCGCAACGAAACCAAGATGCGCGCGAGTGAACCTTGATTAATGCTTGAAC 1569
QY 1567 GTTGCCTATGTAAGCAATTCAGGATGCTTATTAAGCAAGCTTATTAAGCAAGCTTAT 1626
DB 1570 GTTGCCTATGTAAGCAATTCAGGATGCTTATTAAGCAAGCTTATTAAGCAAGCTTAT 1629
QY 1627 ACTGCAAGCGCGCGCAAGCAATTAAGCAAGCTTATTAAGCAAGCTTATTAAGCAAGCTT 1686
DB 1630 ACTGCAAGCGCGCGCAAGCAATTAAGCAAGCTTATTAAGCAAGCTTATTAAGCAAGCTT 1689
QY 1687 ACTACTGCTTGTAACTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1746
DB 1690 ACTACTGCTTGTAACTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1749
QY 1747 ACAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1806
DB 1750 ACAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1809
QY 1807 AACGTAAGAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1866
DB 1810 AACGTAAGAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1869
QY 1867 GCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
DB 1870 GCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1923

RESULT 10
US-10-820-712A-11
; Sequence 11, Application US/10820712A
; Publication No. US2005026804A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION

APPLICANT: Okuda, Mitsuoshi
APPLICANT: Izawa, Yoshihumi
APPLICANT: Kobayashi, Tohru
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-USO
CURRENT APPLICATION NUMBER: US/10/820,712A
PRIORITY FILING DATE: 2004-04-09
PRIORITY APPLICATION NUMBER: 2003-106708
PRIORITY FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.2
SEQ ID NO 11
LENGTH: 1302
TYPE: DNA
ORGANISM: Bacillus sp. KSM-kp9860
FEATURE:
NAME/KEY: CDS (1302)
LOCATION: (1)..
OTHER INFORMATION: protease KP9860
US-10-820-712A-11

Query Match 67.5%; Score 1295.6; DB 8; Length 1302;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1298; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 616 AATGATGCGCCAGAGATTGTCAAAGCGGATGTGGCAGACAGACGCTGATTGAT 675
DB 1 AATGATGCGCCAGAGATTGTCAAAGCGGATGTGGCAGACAGACGCTGATTGAT 60
QY 676 GCACAAAGCCAGATTGTGCGAGTTGCCGATCTGATGATGATGATGATGATGAT 735
DB 61 GCACAAAGCCAGATTGTGCGAGTTGCCGATCTGATGATGATGATGATGATGAT 120
QY 736 TCGATGATGAAAGCTTCCGCGTAAATTAACAGCACTATATGACTGGTCCGAGAT 795
DB 121 TCGATGATGAAAGCTTCCGCGTAAATTAACAGCACTATATGACTGGTCCGAGAT 180
QY 796 AATGGAATGATACGAACGGTATGTAACCCATGTGGCAGGTTCCGATTAAGAAATGCG 855
DB 181 AATGGAATGATACGAACGGTATGTAACCCATGTGGCAGGTTCCGATTAAGAAATGCG 240
QY 856 GCAACGAATTAAGAAATGCGACCTCAAGCGAATCTGTTTTCATTCATCATGATGATGAC 915
DB 241 GCAACGAATTAAGAAATGCGACCTCAAGCGAATCTGTTTTCATTCATCATGATGATGAC 300
QY 916 AGTGTGGGCTTGAAGGCTTGCCTTCAATCTGCAAACTTATTCAGCCAAACATTCAGT 975
DB 301 AGTGTGGGCTTGAAGGCTTGCCTTCAATCTGCAAACTTATTCAGCCAAACATTCAGT 360
QY 976 GCAAGTGCAGAAATTCATCAAACTCTGCGGGGGGAGCGGTGAATGGGGCTTACAGCA 1035
DB 361 GCAAGTGCAGAAATTCATCAAACTCTGCGGGGGGAGCGGTGAATGGGGCTTACAGCA 420
QY 1036 GATTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095
DB 421 GATTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 1096 GGGAAATGAAGCGCGAAGCGGGTACATCACTGCACTGCTGCTGCTGCTGCTGCTGCT 1155
DB 481 GGGAAATGAAGCGCGAAGCGGGTACATCACTGCACTGCTGCTGCTGCTGCTGCTGCT 540
QY 1156 ACAGTCCGCGGCAACGGAACCTGCTCAAGCTTCCGTTCTATGAGATTAATTTTAC 1215
DB 541 ACTGTCCGCGGCAACGGAACCTGCTCAAGCTTCCGTTCTATGAGATTAATTTTAC 600
QY 1216 CACGTTGACAGATTCTTCCCGTGGCCGCAAAAGATGGCGGATCAAGCTGATGTC 1275
DB 601 CACGTTGACAGATTCTTCCCGTGGCCGCAAAAGATGGCGGATCAAGCTGATGTC 660
QY 1276 ATGGCGCAGGAGACATATTTATCAGCAAGATCTTCTTTCGACCCGATTCCTCCTTC 1335
DB 1276 ATGGCGCAGGAGACATATTTATCAGCAAGATCTTCTTTCGACCCGATTCCTCCTTC 1335

DB 661 ATGGCGCAGGAGACATATTTATCAGCAAGATCTTCTTTCGACCCGATTCCTCCTTC 720
QY 1336 TGGCGGAATCATGACAGAAATATGCTTATATGCGTGAAGCTGCATGCAACCGATT 1395
DB 721 TGGCGGAATCATGACAGAAATATGCTTATATGCGTGAAGCTGCATGCAACCGATT 780
QY 1396 GTTCGCGGGAATGTGACAGCTCCGTGAGATTTTGTGAAAATATGAGAAATCACTCCT 1455
DB 781 GTTCGCGGGAATGTGACAGCTCCGTGAGATTTTGTGAAAATATGAGAAATCACTCCT 840
QY 1456 AAGCTTCCCTATTAAGAAAGCACTTTGATGATGATGATGATGATGATGATGATGAT 1515
DB 841 AAGCTTCCCTATTAAGAAAGCACTTTGATGATGATGATGATGATGATGATGATGAT 900
QY 1516 CCGAACGGAACCAAGATGGGGCCGAGTGACCTCGATTAATGTTGAAGCTTCTAT 1575
DB 901 CCGAACGGAACCAAGATGGGGCCGAGTGACCTCGATTAATGTTGAAGCTTCTAT 960
QY 1576 GTGAACGAATCCAGTCCCTATCAACTAGCCAAAAGGACATATACCTTAAGCAACG 1635
DB 961 GTGAACGAATCCAGTCCCTATCAACTAGCCAAAAGGACATATACCTTAAGCAACG 1020
QY 1636 GCGGCGAAGCCATTGAATAATCTCCGTGTATGATGATGATGATGATGATGATGATGAT 1695
DB 1021 GCGGCGAAGCCATTGAATAATCTCCGTGTATGATGATGATGATGATGATGATGATGAT 1080
QY 1696 TCTGTAAACCTGTGCAATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1755
DB 1081 TCTGTAAACCTGTGCAATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1140
QY 1756 GTGCGGAATGACTCTCGACACATTTGACAAATTAAGTGGATGCGCAATTAAGTGA 1815
DB 1141 GTGCGGAATGACTCTCGACACATTTGACAAATTAAGTGGATGCGCAATTAAGTGA 1200
QY 1816 AATGATTTTATTAATTCGCCCCAAGTGAAATATTAATTAATTAATTAATTAATTAAT 1875
DB 1201 AATGATTTTATTAATTCGCCCCAAGTGAAATATTAATTAATTAATTAATTAATTAAT 1260
QY 1876 GTGCGGTTGACCAAAACCTTCTGTTGGCAATTTGTAAC 1917
DB 1261 GTGCGGTTGACCAAAACCTTCTGTTGGCAATTTGTAAC 1302

RESULT 11

US-10-820-714A-12

Sequence 12, Application US/10820714A

Publication No. US20050214922A1

GENERAL INFORMATION:

APPLICANT: KAO CORPORATION

APPLICANT: Okuda, Mitsuoshi

APPLICANT: Kobayashi, Tohru

APPLICANT: Sumitomo, Nobuyuki

APPLICANT: Takimura, Yasushi

TITLE OF INVENTION: ALKALINE PROTEASE

FILE REFERENCE: 251697USO

CURRENT APPLICATION NUMBER: US/10/820,714A

PRIORITY FILING DATE: 2004-04-09

PRIORITY APPLICATION NUMBER: 2003-106709

PRIORITY FILING DATE: 2003-04-10

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn version 3.2

SEQ ID NO 12

LENGTH: 1302

TYPE: DNA

ORGANISM: Bacillus sp. KSM-kp9860

FEATURE:

NAME/KEY: CDS (1302)

LOCATION: (1)..
OTHER INFORMATION: protease KP9860

US-10-820-714A-12

Query Match

67.5%; Score 1295.6; DB 9; Length 1302;

Best Local Similarity 99.7%; Pred. No. 0;
Matches 1298; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 616 AATGATGTCGCGGAGGATATTTGCAAAAGCGGATGTGCAAGAGCACTACGGTTGTAT 675
DB 1 AATGATGTCGCGGAGGATATTTGCAAAAGCGGATGTGCAAGAGCACTACGGTTGTAT 60
QY 676 GGAACAAGCCAGATTTGTGCGGATGTGCAAGATGTGCAAGATGTGCAAGATGTGCAAGAT 735
DB 61 GGAACAAGCCAGATTTGTGCGGATGTGCAAGATGTGCAAGATGTGCAAGATGTGCAAGAT 120
QY 736 TCGATGATGAAAGCTTCCGCGGTAAATTAACGCACTATATGCTACGTGGTGGAGAT 795
DB 121 TCGATGATGAAAGCTTCCGCGGTAAATTAACGCACTATATGCTACGTGGTGGAGAT 180
QY 796 AATGCGAATGATTCGAAGCGGTCTATGCTACGTGGTGGAGATGTGCAAGATGTGCAAGAT 855
DB 181 AATGCGAATGATTCGAAGCGGTCTATGCTACGTGGTGGAGATGTGCAAGATGTGCAAGAT 240
QY 856 GCAACGAATTAAGGATGTGCAAGCGGATCTGAGTTTTCATTCATTCATTCATTCATTCAT 915
DB 241 GCAACGAATTAAGGATGTGCAAGCGGATCTGAGTTTTCATTCATTCATTCATTCATTCAT 300
QY 916 AATGCGGCTTGGAGGCTTGCCTTCATTCGCAAACTTATTCAGCAAGCAATTCAGT 975
DB 301 AATGCGGCTTGGAGGCTTGCCTTCATTCGCAAACTTATTCAGCAAGCAATTCAGT 360
QY 976 GCAAGTTCGCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1035
DB 361 GCAAGTTCGCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 420
QY 1036 GATTCGAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095
DB 421 GATTCGAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 1096 GGAATGAAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1155
DB 481 GGAATGAAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 1156 AATGCGGCTTGGAGGCTTGCCTTCATTCGCAAACTTATTCAGCAAGCAATTCAGT 1215
DB 541 AATGCGGCTTGGAGGCTTGCCTTCATTCGCAAACTTATTCAGCAAGCAATTCAGT 600
QY 1216 GATTCGAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
DB 601 GATTCGAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 1276 AATGCGGCTTGGAGGCTTGCCTTCATTCGCAAACTTATTCAGCAAGCAATTCAGT 1335
DB 661 AATGCGGCTTGGAGGCTTGCCTTCATTCGCAAACTTATTCAGCAAGCAATTCAGT 720
QY 1336 TGGGCGAATTCATGACGCAAAATATGCTATATGAGTGAACGTTCATTCAGCAAGCAATTCAGT 1395
DB 721 TGGGCGAATTCATGACGCAAAATATGCTATATGAGTGAACGTTCATTCAGCAAGCAATTCAGT 780
QY 1396 GTTTCGCGAATTCATGACGCAAAATATGCTATATGAGTGAACGTTCATTCAGCAAGCAATTCAGT 1455
DB 781 GTTTCGCGAATTCATGACGCAAAATATGCTATATGAGTGAACGTTCATTCAGCAAGCAATTCAGT 840
QY 1456 AATGCGGCTTGGAGGCTTGCCTTCATTCGCAAACTTATTCAGCAAGCAATTCAGT 1515
DB 841 AATGCGGCTTGGAGGCTTGCCTTCATTCGCAAACTTATTCAGCAAGCAATTCAGT 900
QY 1516 GGAACAAGCCAGATTTGTGCGGATGTGCAAGATGTGCAAGATGTGCAAGATGTGCAAGAT 1575
DB 901 GGAACAAGCCAGATTTGTGCGGATGTGCAAGATGTGCAAGATGTGCAAGATGTGCAAGAT 960
QY 1576 TCGATGATGAAAGCTTCCGCGGTAAATTAACGCACTATATGCTACGTGGTGGAGAT 1635
DB 961 TCGATGATGAAAGCTTCCGCGGTAAATTAACGCACTATATGCTACGTGGTGGAGAT 1020
QY 1636 GCGGCGAAGCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1695
DB 1695 GCGGCGAAGCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT
```

```
DB 1021 GCGGCGAAGCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1080
QY 1696 TCGATGATGAAAGCTTCCGCGGTAAATTAACGCACTATATGCTACGTGGTGGAGAT 1755
DB 1081 TCGATGATGAAAGCTTCCGCGGTAAATTAACGCACTATATGCTACGTGGTGGAGAT 1140
QY 1756 GCGGCGAAGCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1815
DB 1141 GCGGCGAAGCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1200
QY 1816 AATGATTTTATTAATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1875
DB 1201 AATGATTTTATTAATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1876 GTGCGGTTGACCAAACTTCTGTTGGCAATTTGTAAC 1917
DB 1261 GTGCGGTTGACCAAACTTCTGTTGGCAATTTGTAAC 1302
```

RESULT 12

```
US-10-385-662-1
; Sequence 1, Application US/10385662
; Publication No. US2004002432A1
; GENERAL INFORMATION:
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: SAITO, KAZUHIRO
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: IZAWA, YOSHIKUMI
; APPLICANT: SAKI, KATSUHIISA
; APPLICANT: KOBAYASHI, TOHRU
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: Alkaline protease
; FILE REFERENCE: 2349380S0
; CURRENT APPLICATION NUMBER: US/10/385,662
; CURRENT FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: JP 2002-081428
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: JP 2002-165987
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: JP 2002-304230
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: JP 2002-304231
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-KP43
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1305)
; OTHER INFORMATION:
```

Query Match 52.2%; Score 1002.6; DB 6; Length 1305;
Best Local Similarity 85.5%; Pred. No. 6; 6e-299;
Matches 1116; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

```
QY 616 AATGATGTCGCGGAGGATATTTGCAAAAGCGGATGTGCAAGAGCACTACGGTTGTAT 675
DB 1 AATGATGTCGCGGAGGATATTTGCAAAAGCGGATGTGCAAGAGCACTACGGTTGTAT 60
QY 676 GGAACAAGCCAGATTTGTGCGGATGTGCAAGATGTGCAAGATGTGCAAGATGTGCAAGAT 735
DB 61 GGAACAAGCCAGATTTGTGCGGATGTGCAAGATGTGCAAGATGTGCAAGATGTGCAAGAT 120
QY 736 TCGATGATGAAAGCTTCCGCGGTAAATTAACGCACTATATGCTACGTGGTGGAGAT 795
DB 121 TCGATGATGAAAGCTTCCGCGGTAAATTAACGCACTATATGCTACGTGGTGGAGAT 180
QY 796 AATGCGAATGATTCGAAGCGGTCTATGCTACGTGGTGGAGATGTGCAAGATGTGCAAGAT 855
DB 855 AATGCGAATGATTCGAAGCGGTCTATGCTACGTGGTGGAGATGTGCAAGATGTGCAAGAT
```


Db 101 AATGCAATGATACGATGCTCATGTGATGCGATGCGGCTCCGATATTAGAAACGGC 240
QY 856 GCAACGAATAAAGAAATGCGACCTTCAAGGAAATCTGCTTTTCAATCCATCATGATAGC 915
Db 241 TCCACTAATAAAGAAATGCGCTCCGCGGAAATCTAGCTTCCAAATCTATCATGTAGTAC 300
QY 916 AGTGTGGGCTTGGAGGCTTGCCTTCATCTGCAAAACCTTATTTAGAGCAACATTCAGT 975
Db 301 GGTGGGGGACTTGGAGACTTACCTTGATCTGCAAACTTATTTACGCGCAACATACAGT 360
QY 976 GCAAGTCCGAAATTCATCAAACTCTGCGGGGCGACCGGTGAATGGGGCTTACAGACA 1035
Db 361 GGTGGTCCAGAAATTCATCAAACTCTGCGGAGAGAGAGATGAGGGCTTACAGACA 420
QY 1036 GATTCCAGAAATGTGATGATGATGTAAGGAAATATGATATGACGATTTCTTTCGGGCT 1095
Db 421 GATTCAGAAATGTGATGATGATGTAAGGAAATATGATATGACGATTTCTTTCGGGCT 480
QY 1096 GGGAAATGAAGGCGGAAACGGCGGTACCATGATGACCTGATGCGCTTAAACCGCATTA 1155
Db 481 GGGAAATGAAGGCGGAAACGGCGGAAACCATGATGACCAAGGCAAGCTTAAATGCAATA 540
QY 1156 ACAGTCCGCGGCAACCGGAAACCTGCGTCCAGACTTCGGTTCCATGAGATTAATTAAC 1215
Db 541 ACAGTCCGAGCTACCGGAAACCTCCGCGCAAGCTTTGGGTCTTATGCGGACAAATTAAC 600
QY 1216 CAGCTTGCACAGTTCTCTCCCGTGGCCGCAAAAGATGGCGGAATCAAGCTTATGTC 1275
Db 601 CATGTGGCAAGATTTCTTTCAGTGGACCGAACAAAGATGACGATCAACCGGATGTC 660
QY 1276 ATGGGCGGAGGAGCATCAATTTATCAGGAATCTTCTGTCGACCGGATTCCTTCCTTC 1335
Db 661 ATGGCAACGGGAACTTCACTATCAAGCAAGTCTTCTTTCGACCGGATTCCTTCCTTC 720
QY 1336 TGGGGAATTCATGACAGCAAAATATGCTATATGCTGGAACCTTCATGCGACACCGATT 1395
Db 721 TGGGGAACCAATGACAGTAATATGATATGATGAGGAGAACTCCATGCTTACACGATC 780
QY 1396 GTTGGGGAATGTGACAGCTCCGTGAGCATTTTGTGAAAAATAGAGAAATCACTCTCT 1455
Db 781 GTTGTGGAACGTGCAAGCTCTGTGAGCATTTTGTGAAAAACAGAGCATCAACACA 840
QY 1456 AAGCTTCCTTATGAAAGCAGCTTATGATGAGGAGGCTGATGTTGGATTTGGGTTAT 1515
Db 841 AAGCTTCTCTATTAAGGCGGCACTGATGCGGAGCTGACAGATGGGCTTTGGCTAC 900
QY 1516 CCGAACGGAACCAAGATGGGGCGAGTGAACCTGATTAATCGTTGAACGTTGCTTAT 1575
Db 901 CCGAACGGAACCAAGATGGGGAGAGTGAACATTTGGATTAATCCCTGAAGCTTGGCTAT 960
QY 1576 GTGAACGAATCGAGTGGCTTCAATGAGCCAAAGAGCAATATCCTTTATGCAACG 1635
Db 961 GTGAACGAGTCAAGTCTCTATCAACAGCCAAAGAGCACTTACCTGTTATCTGCTACT 1020
QY 1636 GCGGCGAAGCAATGAAATCTCCGTGATATGATGAGTGGCGCGCAAGCATGCTGCT 1695
Db 1021 GCGGCGAAGCAATGAAATCTCCGTGATATGATGAGTGGCGCGCAAGCATGCTGCT 1080
QY 1696 TCTGTAACTGTGCAATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1755
Db 1081 TCCGTAAAGCTTGTCAATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1140
QY 1756 GTCCGGAATGACTTCTCAGCAACATTTGACAAATCTGGGATGGCGGCAATATGTA 1815
Db 1141 GTAGGAATATGACTTATCTGCGCATATGATTAATCTGGGATGGCGGCAATATGTA 1200
QY 1816 AATGATTTAATTAATGCGCCCAAGTGAATGATTAATCAATGAGTGAAGCAATATAT 1875
Db 1201 AATGATTTAATTAATGCGCCCAAGTGAATGATTAATCAATGAGTGAAGCAATATAT 1260
QY 1876 GTGCGGTTGACCAAAACCTTCTGTTGGCAATTTGAACTAA 1920

Db 1261 GTACCGGTTGACCAAGACCTTCTGTTGGCAATTTGAAATTA 1305
RESULT 13
US-10-820-712A-13
; Sequence 13, Application US/10820712A
; Publication No. US2005026804A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuyoshi
; APPLICANT: Izawa, Yoshihumi
; APPLICANT: Kobayashi, Tooru
; APPLICANT: Koyama, Shingo
; APPLICANT: Sato, Tsuyoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 251701-US0
; CURRENT APPLICATION NUMBER: US/10/820,712A
; CURRENT FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106708
; PRIOR FILING DATE: 2003-04-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-9865
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: Protease 9865
US-10-820-712A-13
Query Match 52.1%; Score 999.6; DB 8; Length 1302;
Best Local Similarity 85.5%; Pred. No. 5.6e-258;
Matches 1113; Conservative 0; Mismatches 189; Indels 0; Gaps 0;
QY 616 AATGATGCGCAGAGGATTTGCAAGCGGATGTCAGACAGACGATCGTTGAT 675
Db 1 AATGATGTCAGCTGGAATTTGCAAGCGGATGTCAGACAGACGATCGTTGAT 60
QY 676 GGAAGGCGAGATTTGTCAGATTTGCCGATTACTGATTTGATATGAGAAAGACAGT 735
Db 61 GGAAGGAGACAGATCTGATGCGGTTGCCGATTCAGGGCTTGTATCAGGTCCGATGACAGT 120
QY 736 TCGATGATGAAGCTTCCGCGGTAAATTAACAGACATATATGCACTGGGTCCGAGAT 795
Db 121 TCGATGATGAAGCTTCCGCGGTAAATTAACAGACATATATGCACTGGGTCCGAGAT 180
QY 796 AATGCAATGATGAAAGCGTCAATGATACCAATGTCAGAGGTTGATTAAGAAATGGC 855
Db 181 AATGCAATGATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 856 GCAACGAATAAAGAAATGCGACCTTCAAGCGAATCTGCTTTTCAATCATCATGATAGC 915
Db 241 TCCACTAATAAAGAAATGCGCTCCGCGGAAATCTAGCTTCCAAATCTATCATGTAGTAC 300
QY 916 AGTGTGGGCTTGGAGGCTTGCCTTCATCTGCAAAACCTTATTTAGAGCAACATTCAGT 975
Db 301 GGTGGGGGACTTGGAGACTTACCTTGATCTGCAAACTTATTTACGCGCAACATACAGT 360
QY 976 GCAAGTCCGAAATTCATCAAACTCTGCGGGGCGACCGGTGAATGGGGCTTACAGACA 1035
Db 361 GGTGGTCCAGAAATTCATCAAACTCTGCGGAGAGAGAGATGAGGGCTTACAGACA 420
QY 1036 GATTCCAGAAATGTGATGATGATGTAAGGAAATATGATATGACGATTTCTTTCGGGCT 1095
Db 421 GATTCAGAAATGTGATGATGATGTAAGGAAATATGATATGACGATTTCTTTCGGGCT 480
QY 1096 GGGAAATGAAGGCGGAAACGGCGGTACCATGATGACCTGATGCGCTTAAACCGCATTA 1155
Db 481 GGGAAATGAAGGCGGAAACGGCGGAAACCATGATGACCAAGGCAAGCTTAAATGCAATA 540
QY 1156 ACAGTCCGCGGCAACCGGAAACCTGCGTCCAGACTTCGGTTCCATGAGATTAATTAAC 1215


```

Db      541 AAGAGTGGAGGTACGAGAAAACCTCCGCCAAGCTTGCGGTCTATGCGAGCAATATCAAC 600
QY      1216 CACGTTGCAAGTCTCTTCCGTGCGCCGACAAAGATGGCGGAATCAAGCTGATGTC 1275
Db      601 CATGTGGCAAGTCTCTTCACTGAGCCGACAAAGATGAGCGGATCAACCGGATGTC 660
QY      1276 ATGGCGCCAGGGACATACATTTTATGAGAAATCTTCTCTTGACCCGATTCCTCTTC 1335
Db      661 ATGGCACCGGGAACGTTCACTATACAGAAATCTTCTCTTGACCCGATTCCTCTTC 720
QY      1336 TGGCGCAATCATGACAGCAATATGCTTATGAGTGAACGTCCATGCAACACCGATT 1395
Db      721 TGGGGGAAACATGACAGTAAATATGATATCATGGGTGGAACGTCCATGCTACACCGATC 780
QY      1396 GTTGGCGGGAATGTTGCAAGCTCCGTGACATTTTGTGAAAATAGAGAACTACTCT 1455
Db      781 GTTGGCTGGAACGTGGCAAGCTTCGTGAGCAATTTTGTGAAAACAGAGGCATCACACCA 840
QY      1456 AAGCCTTCCTATTTGAAAGCAGCTTGTGATTCAGGTCGTGATGTTGGATTGGATTAT 1515
Db      841 AAGCCTTCCTATTTGAAAGCAGCTGATGATTCGCGGTGAGCAAGCAATCGGCTTGCTAC 900
QY      1516 CCGAACGGAAACCAAGATGGGGCCGAGTGAACCTGATTAATCGTTGAACGTTGCTTAT 1575
Db      901 CCGAACGGTAACCAAGATGGGAGAGATGACATTTGATTAATCCTTAACGTTGCTTAT 960
QY      1576 GTGAAAGATTCAGTGCCTTATCACTAGCCAAAGACGACATATATCTTTACTGCAACG 1635
Db      961 GTGAAAGATTCAGTGCCTTATCACTAGCCAAAGACGACATATATCTTTACTGCTAT 1020
QY      1636 GGGGGCAAGCCATTGAAAATCTCCCTGTGATGTCGGAATGCCCTGCAAGCACTACTGCT 1695
Db      1021 GCGGCAAGCCCTTGAAGATCTCCCTGTGATGTCGGAATGCCCTGCAAGCACTACTGCT 1080
QY      1696 TCTGTAACTGTCATATGATTTGATTTGTCATTTAGACAGCAACAAAGCAAGATAT 1755
Db      1081 TCCGTAACTGTCATATGATTTGATTTGTCATTTAGACAGCAACAAAGCAAGATAT 1140
QY      1756 GTGGGAATGACTTCTGACGACCATTTGACAAATATGAGATGGCGCAATATACGTAGA 1815
Db      1141 GTTGAATATGACTTACTTCCCATATCAATATATATCTGGGATGGCGCAATATACGTAGA 1200
QY      1816 AATGATTTATTAATTTCCGCCCAAGTGAACATATACATTTGAGGTGCAAGCATATAT 1875
Db      1201 AATGATTTATTAATTTCCGCCCAAGTGAACATATACATTTGAGGTGCAAGCATATAT 1260
QY      1876 GTGGCGTTGACGACAAACCTTCTGTTGCAATTTGGAAC 1917
Db      1261 GTACCGGTTGACGACAAACCTTCTGTTGCAATTTGGAAC 1302

RESULT 14
US-10-820-714A-14
; Sequence 14, Application US/10820714A
; Publication No. US20050214922A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuyoshi
; APPLICANT: Kobayashi, Tohru
; APPLICANT: Sumitomo, Nobuyuki
; APPLICANT: Takimura, Yasuohi
; APPLICANT: Sato, Tsuyoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 251697USO
; CURRENT APPLICATION NUMBER: US/10/820,714A
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106709
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1302
```

```

; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-9865
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: Protease 9865
US-10-820-714A-14

Query Match      52.1%; Score 999.6; DB 9; Length 1302;
Best Local Similarity 85.5%; Pred. No. 5.6e-298;
Matches 1113; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

QY      616 AATGATGTGCGCAAGTATTTGTCAAGCGGATGTGGCAACAGCAAGCTTATGAT 675
Db      1 AATGATGTGCAAGTGAATTTGCAAGCGGATGTGGCAAGCAAGCTTATGAT 60
QY      676 GGAAGGCGCAGATTTGTGCAAGTTCGCAATATGATTTGATGATGATGATGATGATGAT 735
Db      61 GGAAGGCAAGACATGTGACGCTTCCGATACAGAGGCTTGAATACAGGTGCAATGACAT 120
QY      736 TCGATGCAATGAAGCTTCCGCGGTAAATATACAGCACTATATGCACTGGGTGCAAGAT 795
Db      121 TCGATGCAATGAAGCTTCCGCGGTAAATATACAGCACTATATGCACTGGGTGCAAGAT 180
QY      796 AATGCGAATGATACGAAAGGTATGATCCCATGTGGCAGGTTCGGTATTAGAAATGCG 855
Db      181 AATGCGCAATGATACGAAAGGTATGATCCCATGTGGCAGGTTCGGTATTAGAAATGCG 240
QY      856 GCAAGCAATTAAGAAATGGCACTCAAGCGAATCTGTTTCAATCCATCAATGATAGC 915
Db      241 TCCATCAATTAAGAAATGGCACTCAAGCGAATCTGTTTCAATCCATCAATGATAGC 300
QY      916 AGTGTGGGCTTGAAGCTTGTGCTTCAATCTGCAACCTTATTCAGCCAGCATTCAGT 975
Db      301 GGTGGGGGACTTGAAGCTTGTGCTTCAATCTGCAACCTTATTCAGCCAGCATTCAGT 360
QY      976 GGAAGTGCAGAAATTCATACAACTCTGGGGGGGAGCGGTAAATGGGGCTTACACAGCA 1035
Db      361 GCTGTGTCAGAAATTCATACAACTCTGGGGGGGAGCGGTAAATGGGGCTTACACAGCA 420
QY      1036 GATTCAGAAATGTGATGATGATATGAAAGAAATGATATGATGATGATGATGATGAT 1095
Db      421 GATTCAGAAATGTGATGATGATATGAAAGAAATGATATGATGATGATGATGATGAT 480
QY      1096 GGAATGAAAGCGCCGAACGGCGGTACCATGATGACCTGTGATCGGCTTAACACCGCAT 1155
Db      481 GGAATGAAAGCGCCGAACGGCGGTACCATGATGACCTGTGATCGGCTTAACACCGCAT 540
QY      1156 ACAGTGGCGCAACCGAAACCTGCTGCAAGCTTGGGTTCTATGCAATATATTAAC 1215
Db      541 ACAGTGGAGCTACCGAAACCTCGCCCAAGCTTCGGGTCTTATGCGGACAAATATCAAC 600
QY      1216 CACGTGCAAGTCTCTCCCGTGGCCCGCAAAAGATGGCGGAATCAAGCTGATATGTC 1275
Db      601 CATGTGGCAAGTCTCTTCACTGAGCCGACAAAGATGAGCGGATCAACCGGATGTC 660
QY      1276 ATGGCGCCAGGACATACATTTTATCAGCAAGATCTTCTTTCGACCCGATTCCTCTTC 1335
Db      661 ATGGCACCGGGAACGTTCACTATACAGAAATCTTCTTTCGACCCGATTCCTCTTC 720
QY      1336 TGGCGCAATCATGACAGCAATATGCTTATGAGTGAACGTCCATGCAACACCGATT 1395
Db      721 TGGCGCAACATGACAGTAAATATGCAATATGAGTGAACGTCCATGCAACACCGATC 780
QY      1396 GTTGGCGGGAATGTTGCAAGCTCCGTGACATTTTGTGAAAATAGAGAAATCACTCT 1455
Db      781 GTTGGCTGGAACGTGGCAAGCTTCGTGAGCAATTTTGTGAAAACAGAGGCATCACACCA 840
QY      1456 AAGCCTTCCTATTTGAAAGCAGCTTGTGATTCAGGTCGTGATGTTGGATTGGATTAT 1515
Db      841 AAGCCTTCCTATTTGAAAGCAGCTGATTTCCGCGGTGAGCAAGCATGCGCTTGCTGAC 900
QY      1516 CCGAACGGAAACCAAGATGGGGCCGAGTGAACCTTGATTAATCGTTGAACGTTGCTTAT 1575
```


Db 1261 GTTCCACAGGGCCGCGCTTTCTTGGCTATTGTGAC 1302

Search completed: April 8, 2006, 07:34:32
Job time : 1381.42 secs

This Page Blank (usp10)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 03:01:47 ; Search time 896.3 Seconds
(without alignments)
8569.753 Million cell updates/sec

Title: US-10-784-870-3

Perfect score: 1920

Sequence: 1 atgagaagaagaagtgctt.....cgttcgaatcgtgaactaa 1920

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New.*
1: /SIDS5/ptocdata/1/pubpna/US08_NEW_PUB.seq.*
2: /SIDS5/ptocdata/1/pubpna/US06_NEW_PUB.seq.*
3: /SIDS5/ptocdata/1/pubpna/US07_NEW_PUB.seq.*
4: /SIDS5/ptocdata/1/pubpna/PCT_NEW_PUB.seq.*
5: /SIDS5/ptocdata/1/pubpna/US09_NEW_PUB.seq.*
6: /SIDS5/ptocdata/1/pubpna/US09_NEW_PUB.seq.*
7: /SIDS5/ptocdata/1/pubpna/US10_NEW_PUB.seq.*
8: /SIDS5/ptocdata/1/pubpna/US10_NEW_PUB.seq.*
9: /SIDS5/ptocdata/1/pubpna/US10_NEW_PUB.seq.*
10: /SIDS5/ptocdata/1/pubpna/US11_NEW_PUB.seq.*
11: /SIDS5/ptocdata/1/pubpna/US11_NEW_PUB.seq.*
12: /SIDS5/ptocdata/1/pubpna/US11_NEW_PUB.seq.*
13: /SIDS5/ptocdata/1/pubpna/US11_NEW_PUB.seq.*
14: /SIDS5/ptocdata/1/pubpna/US11_NEW_PUB.seq.*
15: /SIDS5/ptocdata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	1.9	972	US-10-301-480-95183	Sequence 95183, A
2	37	1.9	972	US-10-301-480-708592	Sequence 708592, A
3	36.6	1.9	622	US-09-925-065A-613029	Sequence 613029, A
4	36.2	1.9	600	US-10-301-480-78689	Sequence 78689, A
5	36.2	1.9	620	US-10-301-480-692028	Sequence 692028, A
6	36	1.9	628	US-09-925-065A-600872	Sequence 600872, A
7	36	1.9	628	US-09-925-065A-600873	Sequence 600873, A
8	36	1.9	172649	US-11-121-086-36	Sequence 36, App1
9	35.8	1.9	549	US-09-925-065A-609963	Sequence 609963, A
10	35.8	1.9	580	US-10-301-480-70454	Sequence 70454, A
11	35.8	1.9	580	US-10-301-480-70456	Sequence 70456, A
12	35.8	1.9	580	US-10-301-480-683863	Sequence 683863, A
13	35.8	1.9	580	US-10-301-480-683865	Sequence 683865, A
14	35.8	1.9	650	US-09-925-065A-822653	Sequence 822653, A
15	35.6	1.9	431	US-09-925-065A-600874	Sequence 600874, A
16	35.6	1.9	753	US-10-932-182A-4129	Sequence 4129, App
17	35.6	1.9	753	US-10-932-182A-4129	Sequence 4129, App
18	35.6	1.9	4656	US-10-932-182A-3130	Sequence 3130, App

19	35.6	1.9	4656	US-10-932-182A-3130	Sequence 3130, App
20	35.4	1.8	580	US-10-301-480-70455	Sequence 70455, A
21	35.4	1.8	580	US-10-301-480-683864	Sequence 683864, A
22	35.4	1.8	1175	US-11-102-026A-111	Sequence 111, App
23	35.4	1.8	1215	US-11-102-026A-187	Sequence 187, App
24	35	1.8	1758	US-10-821-234-637	Sequence 637, App
25	35	1.8	611587	US-11-117-187-309	Sequence 209, App
26	34.6	1.8	390183	US-10-330-773-937	Sequence 937, App
27	34.4	1.8	709	US-09-925-065A-866816	Sequence 866816, A
28	34.4	1.8	995	US-10-301-480-564083	Sequence 564083, A
29	34.4	1.8	995	US-10-301-480-1177492	Sequence 1177492, A
30	34.4	1.8	159146	US-11-121-086-49	Sequence 49, App1
31	34.2	1.8	452	US-09-925-065A-199670	Sequence 199670, A
32	34.2	1.8	455	US-10-301-480-288350	Sequence 288350, A
33	34.2	1.8	455	US-10-301-480-901759	Sequence 901759, A
34	34.2	1.8	5617	US-11-136-527-2014	Sequence 2014, App
35	34	1.8	515	US-09-925-065A-701445	Sequence 701445, A
36	34	1.8	722	US-09-925-065A-85259	Sequence 85259, A
37	34	1.8	722	US-09-925-065A-85260	Sequence 85260, A
38	34	1.8	722	US-10-301-480-186499	Sequence 186499, A
39	34	1.8	722	US-10-301-480-186500	Sequence 186500, A
40	34	1.8	722	US-10-301-480-799908	Sequence 799908, A
41	34	1.8	722	US-10-301-480-799909	Sequence 799909, A
42	34	1.8	998	US-10-301-480-562054	Sequence 562054, A
43	34	1.8	998	US-10-301-480-1175463	Sequence 1175463, A
44	34	1.8	1051	US-09-925-065A-64075	Sequence 64075, A
45	34	1.8	1051	US-10-301-480-165313	Sequence 165313, A

ALIGNMENTS

RESULT 1
US-10-301-480-95183
Sequence 95183, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 95183
LENGTH: 972
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-95183

Query Match 1.9%; Score 37; DB 9; Length 972;
Best Local Similarity 51.5%; Pred. No. 4.4;
Matches 85; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

1167 AACGAAACCTGCTCAAGTTCCTTCATGACATATTAACCAAGTTCGACA 1226
525 AATGTCGAAAGGCACTTAATTAAGTTCATATTAATCACTTAGAGACTTGA 584
1227 GTTCTCCCGCGGCGCAAAAGATGGGCAATCAAGCTGATGTCATGCGCCAG 1286
565 ATTCTGATGCCAGCCACATCATATTAATTAACCTCTTAAGAGGAGATCCAGA 644
1287 GACATACATTTTATCAGCAAGATCTTCTTCGACCCGATTCCTC 1331
645 CATGATAGTTTCTTAAGTGTCCCTAACCCAGGATTCAC 689

RESULT 2

US-10-301-480-708592
; Sequence 708592, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1225818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 708592
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-708592

Query Match 1.9%; Score 37; DB 10; Length 972;
Best Local Similarity 51.5%; Pred. No. 4.4;
Matches 85; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Oy 1167 AACGAAACCTGCTCCAGCTTGGTTCCTATGCAATATTAATACAGCTTGACA 1226
Db 525 AATGTCAGAGGCACTTTAACTTAGTCATATTAATCACTTAGAGAACTTTGAAA 584
Oy 1227 GTTCTCTCCGCGCCGCAAAAGATGGCGAATCAAGCCTGATGTCATGGCCGAG 1286
Db 585 ATTCTGATGCCAGGCCACATCATATGAAATTAACCTCTGAAGAGAGATCCAGA 644
Oy 1287 GACATACATTTTATCAGCAAGATCTTCTCTGACCCCATTCCTC 1331
Db 645 CATTGATAGTTTCTTAAGTGTCCCTACCCCAAGGTGATTCAC 689

RESULT 3
US-09-925-065A-613029/c
; Sequence 613029, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 613029
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-613029

Query Match 1.9%; Score 36.6; DB 6; Length 622;
Best Local Similarity 46.6%; Pred. No. 4.6;
Matches 117; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
Oy 109 TTAAAGAAATTCAGCAAGCAACGATGTCAGTGGTTTCTCAAAACAGCAAAAGGT 168

Db 441 TTTTTCGAATCABATTTTGAATATGTATATGAAGCTTTAAACATTCATATCAAT 382
Oy 169 GCGGCTGATTTCTGCGAGACTGGAATATGTGAACCTTTAAAGATTCCTAAAGAA 228
Db 381 CATTGACGATTTCACTTTTGGTCCCAACATTTACATACATAGATTTGTGTATAT 322
Oy 229 CTTGAACAGTACCGGCAATATTAATCTCATATTTGCCAATTCAGCCCATTTTA 288
Db 321 GCAACTAATGTATTTCCAAATATGTTCTTATGTGAGATTTAGCAAGAGCTTTGATA 262
Oy 289 GAAGAAACAAACAGAACTAGAGACAACTGAGCAAAAGATTCGACTACATCCTGAT 348
Db 261 ATTCTACAAATGAGAACTAAATATATATGCAATATATGTATATTAATTTTAT 202
Oy 349 TATGATATAT 359
Db 201 AATATGTAT 191

RESULT 4
US-10-301-480-78689/c
; Sequence 78689, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78689
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-78689

Query Match 1.9%; Score 36.2; DB 9; Length 600;
Best Local Similarity 54.2%; Pred. No. 5.9;
Matches 71; Conservative 1; Mismatches 59; Indels 0; Gaps 0;

Oy 177 ATTTCTGAGAGTCTGAATATGAAATCTTTAAAGATTTGCTAAAGAACTTGAAAC 236
Db 420 ATATAGGGGAAATGTTAAGAAATTTAAACATTTAAATTAAGTAAAGAAAT 361
Oy 237 AGTACCGCAATATTAATCACTCATATTTGCCAATTCATAGCCCATTTTGAAGAAC 296
Db 360 TTAAACAGCAAAATATATGACAGAAATTAATTTKAAATTAATCACTTATATGAAGC 301
Oy 297 AAAACAGAGC 307
Db 300 ACTTAAATATC 290

RESULT 5
US-10-301-480-692098/c
; Sequence 692098, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09

PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 692098
LENGTH: 600
TYPE: DNA
ORGANISM: Homo sapiens
US-10-301-480-692098

Query Match 1.9%; Score 36.2; DB 10; Length 600;
Best Local Similarity 54.2%; Pred. No. 5.9;
Matches 71; Conservative 1; Mismatches 59; Indels 0; Gaps 0;

QY 177 ATTCTGTGAGTCTGAAAATGTAACCTTTAAAGATGCTAAAGAACTTGAAC 236
DB 420 AATATGGGGGAAAATGTTAGAAAATTAACAAATTTAAATTAAGTAAAGCAAAAT 361
QY 237 AGTACCGGCAATTAATTAATCTCATTTGTCATTCATTCAGGCCCATTTTGAAGAAG 296
DB 360 TTAACGCGCAATTAATTAAGCAAGAAAATTAATTAATTAATTAATTAATTAAGAGC 301
QY 297 AAAACGAGAGC 307
DB 300 ACTTAATAATC 290

RESULT 6

US-09-925-065A-600872
Sequence 600872, Application US/09925065A
Publication No. US20040181048A1

GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 600872
LENGTH: 628
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-600872

Query Match 1.9%; Score 36; DB 6; Length 628;
Best Local Similarity 55.6%; Pred. No. 7;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 203 AACTCTTAAGAGTGTCTAAGAACTGAAACGTAACCGCAATATAATCCATA 262
DB 439 AACATTTCAAGGTTCTTAAGAGTCTCAGCAGGCTTGATGTAATTAACCAATT 498
QY 263 TTGTCCAAATTCATGCCCCCATTTTGAAGAAACAAACAGAGCTAGACAACTGGAG 322
DB 499 TTAACCACTCATATACCTTTTATTAAGAAAAAAGAAAAAAGAACTGTGGGTA 558
QY 323 CAAA 326
DB 559 CAAA 562

RESULT 7

US-09-925-065A-600873
Sequence 600873, Application US/09925065A
Publication No. US20040181048A1

GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 600873
LENGTH: 628
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-600873

Query Match 1.9%; Score 36; DB 6; Length 628;
Best Local Similarity 55.6%; Pred. No. 7;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 203 AACTCTTAAGAGTGTCTAAGAACTGAAACGTAACCGCAATATAATCCATA 262
DB 439 AACATTTCAAGGTTCTTAAGAGTCTCAGCAGGCTTGATGTAATTAACCAATT 498
QY 263 TTGTCCAAATTCATGCCCCCATTTTGAAGAAACAAACAGAGCTAGACAACTGGAG 322
DB 499 TTAACCACTCATATACCTTTTATTAAGAAAAAAGAAAAAAGAACTGTGGGTA 558
QY 323 CAAA 326
DB 559 CAAA 562

RESULT 8

US-11-121-086-36
Sequence 36, Application US/11121086
Publication No. US20050266459A1

GENERAL INFORMATION:
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138,6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 36
LENGTH: 172649
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-36

Query Match 1.9%; Score 36; DB 14; Length 172649;
Best Local Similarity 67.1%; Pred. No. 1.5e+02;
Matches 51; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 192 TGAAGATGGAACCTTTTAAGAGTGTCTAAGAACTGGAACAGTAACCGCAATTA 251

Db 141108 TGAAGATGGCACTATATGTGGAATGTAAAGAAAGTTGAAGTCACTAGAGACAGAGAA 141167

Qy 252 TAAACTCCATATGTC 267

Db 141168 TAAATGTTATTC 141183

RESULT 9

US-09-925-065A-609963/c
; Sequence 609963, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 609963
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-609963

Query Match 1.9%; Score 35.8; DB 6; Length 549;

Best Local Similarity 63.2%; Pred. No. 7.4;
Matches 55; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 194 AAAATGTGAACCTCTTAAAGATTGCTAAAGAACTTGAAACAGTACCGGCAATATATA 253

Db 87 AAAATGAGAGATTTTTAAATATATAGCCACAAAATTTGATTAAGAAAAAAATTT 28

Qy 254 AACTCCATATGTCATATCAATGACC 280

Db 27 AAGATCTTTTAAATCTTTTCC 1

RESULT 10

US-10-301-480-70454/c
; Sequence 70454, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301.480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70454
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-70454

Query Match 1.9%; Score 35.8; DB 9; Length 580;

Best Local Similarity 48.7%; Pred. No. 7.7;
Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

Qy 634 ATTGTCAAGCGGATGTGGACAGAGCACTACGTTTGTATGACAAGGCCAGATTGTC 693

Db 374 ATGTTCTTGAGAGCTGAAGCAGCTAGTTACTTGGGGTTAAACATGGGGTCCGCTG 315

Qy 694 GCATTCGCCATCTGATTTGATATACAGAAAGAACAGCACTTCGATGCATGAAGCCCTTC 753

Db 314 GCACATGCCAGAACATTTAGACACAGAAACACACAGAGATTTTAAAGAGAGGCTCTTGAAAC 255

Qy 754 CGCGTAAATTAACAGCACTATATGCACTGGGTGCGACGAATTAATGCAATGATACGAAC 813

Db 254 ATGGGTAGAGAGAGAGAAAGAAAGAACTGCTTCTTAAGAGAAAGGGTCTTGAAAC 195

Qy 814 GGTGATGTACCCATGTGG 832

Db 194 GGAAGAAAGACCATCTGG 176

RESULT 11

US-10-301-480-70456/c
; Sequence 70456, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301.480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70456
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-70456

Query Match 1.9%; Score 35.8; DB 9; Length 580;

Best Local Similarity 48.7%; Pred. No. 7.7;
Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

Qy 634 ATTGTCAAGCGGATGTGGACAGAGCACTACGTTTGTATGACAAGGCCAGATTGTC 693

Db 374 ATGTTCTTGAGAGCTGAAGCAGCTAGTTACTTGGGGTTAAACATGGGGTCCGCTG 315

Qy 694 GCATTCGCCATCTGATTTGATATACAGAAAGAACAGCACTTCGATGCATGAAGCCCTTC 753

Db 314 GCACATGCCAGAACATTTAGACACAGAAACACACAGAGATTTTAAAGAGAGGCTCTTGAAAC 255

Qy 754 CGCGTAAATTAACAGCACTATATGCACTGGGTGCGACGAATTAATGCAATGATACGAAC 813

Db 254 ATGGGTAGAGAGAGAGAAAGAAAGAACTGCTTCTTAAGAGAAAGGGTCTTGAAAC 195

Qy 814 GGTGATGTACCCATGTGG 832

Db 194 GGAAGAAAGACCATCTGG 176

RESULT 12

US-10-301-480-683863/c
; Sequence 683863, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137

```
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 683863
/ LENGTH: 580
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-683863
```

```
Query Match      1.9%; Score 35.8; DB 10; Length 580;
Best Local Similarity 48.7%; Pred. No. 7.7;
Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
```

```
Qy 634 ATTGCAAGCGGATGTGGCAGACGACTACGGTTTGATGACAAAGCCAGATTGTC 693
Db 374 ATGTTCTTGAGAGAGCTGAAGCAGCTAGTTACTTGGGGTATTAACATGGGGTCCGCTG 315
Qy 634 GCAAGTTCGCGATTCGATGTGATACAGAGAAAGCAGATTCGATGAAAGCTTC 753
Db 314 GCACATGCCAGACATTTAGGACACAGAAACACAGAGATTTTAGAGTGGAGGTTTA 255
Qy 754 CGCGGTAAATTAACGCACTATATGACCTGGTGGACGATTAATGCCAATGATGCAAC 813
Db 254 ATGGGTAGAGAGAGAGAGAAAGAACTGCTTCTTAAGAGAGAGGGGCTTGGAAAC 195
Qy 814 GGTCAATGTAACCATGTGG 832
Db 194 GGAAGAAAAAGACATCTGG 176
```

```
RESULT 13
US-10-301-480-683865/c
/ Sequence 683865, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827,137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 683865
/ LENGTH: 580
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-683865
```

```
Query Match      1.9%; Score 35.8; DB 10; Length 580;
Best Local Similarity 48.7%; Pred. No. 7.7;
Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
```

```
Qy 634 ATTGCAAGCGGATGTGGCAGACGACTACGGTTTGATGACAAAGCCAGATTGTC 693
Db 374 ATGTTCTTGAGAGAGCTGAAGCAGCTAGTTACTTGGGGTATTAACATGGGGTCCGCTG 315
Qy 634 GCAAGTTCGCGATTCGATGTGATACAGAGAAAGCAGATTCGATGAAAGCTTC 753
Db 314 GCACATGCCAGACATTTAGGACACAGAAACACAGAGATTTTAGAGTGGAGGTTTA 255
Qy 754 CGCGGTAAATTAACGCACTATATGACCTGGTGGACGATTAATGCCAATGATGCAAC 813
Db 254 ATGGGTAGAGAGAGAGAGAAAGAACTGCTTCTTAAGAGAGAGGGGCTTGGAAAC 195
```

```
Qy 814 GGTCAATGTAACCATGTGG 832
Db 194 GGAAGAAAAAGACATCTGG 176
```

```
RESULT 14
US-09-925-065A-822653
/ Sequence 822653, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827,135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 822653
/ LENGTH: 650
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-822653
```

```
Query Match      1.9%; Score 35.8; DB 6; Length 650;
Best Local Similarity 57.7%; Pred. No. 8.2;
Matches 64; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

```
Qy 1745 GAACAAGATATGTGCGGATGATGATCTTCAGCACCATTGACAAATTAATGATGCGCCGCA 1804
Db 146 GAATATGATATTTAGTGAATGTTTCATGATACCAACAAACAAACCTATGATTAATCTGCA 205
Qy 1805 ATAAGTGAATAATGATATTTATTAATTCGCCCAAAAGTGAACATATACCA 1855
Db 206 CAATATTAATAAGTTATTTGCAATTTATCCAAATTTTAAGCATACCA 256
```

```
RESULT 15
US-09-925-065A-600874
/ Sequence 600874, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827,135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 600874
/ LENGTH: 431
```

TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-600874

Query Match 1.9%; Score 35.6; DB 6; Length 431;
Best Local Similarity 54.8%; Pred. No. 7.5;
Matches 68; Conservative 1; Mismatches 55; Indels 0; Gaps 0;

OY	203	AACTCTTAAAGGATTGCTAAAGAACTTGAACAGTACCGGCAATATTAATTAACCTCAT	262
Db	109	AACATTCTACAAAGCTTCTTAAAGTCTCAGCAGGCTTGAGTTGATTAATGACCAATT	168
OY	263	TTGTCGAATTCATGCCCCCATTTTGAAGAAACAAACAGAGCTAGAGCAACTGAG	322
Db	169	TTAACCACTCATGACCTTTTCTTTTAAAAAAGAAAAAAGAACTGTGGCTA	228
OY	323	CAAA	326
Db	229	CAAA	232

Search completed: April 8, 2006, 03:55:10
Job time : 898.3 secs

Y

Y

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search **us-10-784-870-4.p2n.rge**

Run on: April 8, 2006, 03:02:08 ; Search time 7456 Seconds
(without alignments)
4871.642 Million cell updates/sec

Title: US-10-784-870-4
Perfect score: 3267
Sequence: 1 MRKKVFLSVLSAAAILSTV.....EVQAVNPVGQNPSTLATVN 639

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-Model=frame+g2n.model -Dev=xlp
-O=/abs/ABSSWB_spool/US10784870/runat_07042006_143415_27459/asp_query.fasta_1
-DB=GenBdb1 -Qfmt=fastrap -Suffix=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human0.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCKL
-OUTMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=AB803P
-USER=US10784870 @CGN_1_1_7415 @runat_07042006_143415_27459 -NCPU=6 -ICPU=3
-NO MMAP -MSG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
GenBdb1:
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hcg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Length	DB ID	Description
1	3267	100.0	1920	1 AB046403 Bacillus
2	3267	100.0	1920	6 AR368116 Sequence
3	3267	100.0	1920	6 AR562464 Sequence

4	3072	94.0	1923	6 AR368118 Sequence
5	3072	94.0	1923	6 AR562466 Sequence
6	3066	93.8	1923	1 AB051423 Bacillus
7	3066	93.8	1923	1 CQ888285 Sequence
8	3066	93.8	1923	6 CQ889309 Sequence
9	3066	93.8	1923	6 AR368117 Sequence
10	3066	93.8	1923	6 AR562465 Sequence
11	3053	93.4	1923	1 AB084155 Bacillus
12	2799	85.7	3003	6 AR069954 Sequence
13	2799	85.7	3003	6 BD062155
14	2500.5	76.5	2218	6 E03808
15	2183	66.8	1305	6 AX839476
16	2122	65.0	1302	6 AB046406
17	2002.5	61.3	1299	1 AB046405
18	1998.5	61.2	1299	1 AB046402
19	1991.5	61.0	1299	1 AB046404
20	816	25.0	346148	1 BX842652
21	680	20.8	60006	3 AF268611
22	655	20.0	110000	1 AE017180_35
23	566	17.3	5890	2 DD060086
24	541	16.6	5772	2 AF466309
25	516.5	15.8	6115	2 DD020432
26	511	15.6	5259	2 AF263455
27	477	14.6	1977	2 AR201152
28	477	14.6	1977	6 AR578779
29	477	14.6	1977	6 AR633917
30	443	13.6	1962	6 AR201155
31	443	13.6	1962	6 AR578782
32	443	13.6	1962	6 AR633946
33	443	13.6	9125	6 AR619564
34	443	13.6	12452	1 AE010265
35	435	13.3	14376	1 AE013049
36	435	13.3	110000	1 AF006878_14
37	432	13.2	1977	6 AR633919
38	431	13.2	14351	2 AC096673
39	417.5	12.8	1236	6 AR201146
40	417.5	12.8	1236	6 AR578773
41	417.5	12.8	1566	6 AR633918
42	402.5	12.3	303450	1 SC0939130
43	401.5	12.3	110000	1 BA000028_24
44	376	11.5	110000	1 CP000099_19
45	376	11.5	110000	1 CP000099_20

ALIGNMENTS

RESULT 1	AB046403	1920 bp	DNA	linear	BCT 10-MAY-2002
LOCUS	Bacillus sp. 9860	PROA gene for protease, complete cds.			
DEFINITION	AB046403				
ACCESSION	AB046403.2	GI:20521152			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Bacillus sp. 9860				
REFERENCE	Bacillus sp. 9860				
AUTHORS	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
TITLE	1				
	Saeki, K., Okuda, M., Hatada, Y., Kobayashi, T., Ito, S., Takami, H. and				
	Horikoshi, K.				
	Novel oxidatively stable subtilisin-like serine proteases from				
	alkaliphilic Bacillus spp.: enzymatic properties, sequences, and				
	evolutionary relationships				
JOURNAL	Biochem. Biophys. Res. Commun. 279 (2), 313-319 (2000)				
PUBMED	11118284				
REFERENCE	2 (bases 1 to 1920)				
AUTHORS	Saeki, K.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2000) Katsuhisa Saeki, Kao Corporation,				
	Biological Science Laboratory, Ichikaimachi Akabane 2606, Hoga,				
	Tochigi 321-3497, Japan (E-mail:387185@kasanet.kao.co.jp,				
	Tel:81-265-68-7400, Fax:81-265-68-7403)				
COMMENT	On May 9, 2002 this sequence version replaced gi:12381938.				
FEATURES	Location/Qualifiers				

```
source
1. 1920
/organism="Bacillus sp. 9860"
/mol_type="genomic DNA"
/strain="9860"
/db_xref="taxon:133778"
1. 1920
/gene="PROA"
1. 1920
/gene="PROA"
/codon_start=1
/codon_table=1
/product="protease"
/protein_id="BAB2126.2"
/db_xref="gi:20521153"
/cranlat="MRKKVFLSVLSAAILSTVALNNSAGDARTEDLPKGIQTTF
DVGPSKOROTGAAPFLVESENVKLKGLKLETPVANKLIVQFNPLIESTKOK
LETGAKLIDYIPIYALIVEYEDVQSVRSIEHVESYEPILPKIDPOLFTGAST
LVKALADTKONNKEVOLRGIEIAQVYASNDVHYITAKPEYKMNMDVARGIVKADVA
OSSYGLVGGQIIVADTGLDTRGNDSSMHEAFRGKITALVALGRTNANDHGTG
VAGSVLNGATRKMAPOANTVFGSINDSSGGLPSNLQTLPSQAFSAGARLHTNS
WGAAVNGAYTDSRNVDYVRKNDMTILFAAGNERPNGTISAGTAKNALIVGATEN
LRSPGSYADNIHVAOPSSRGPTDGRKIKDVMAPGTYIISASSIALPDSFVAMND
SKVAYMGTSMAITPVGAVNAQLREHFVKNRGITPKPSILKALILAGADVLGYPNG
VTLVNDLIVITAPNGTRYVGNDFSAFPDNNMDGRNVENVFINSQSGTITIEQAY
NPVGPQNFSLATYN"
```

ORIGIN

Alignment Scores:

```
Pred. No.: 3,14e-169 Length: 1920
Score: 3267.00 Matches: 639
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0
```

US-10-784-870-4 (1-639) x AB046403 (1-1920)

```
QY 1 MetArglyValValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrVal 20
Db 1 ATGAGAAAGAAAGAGTGCTTTTATCTGTTTATCTGATCGACGCAATCTGCGACTGT 60
QY 21 AlaLeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheValGly 40
Db 61 GCATTAAACAAATCCCTCGGCTGTGATGACAGACTTTTGATCTGGATTTTAAAGAAAT 120
QY 41 GlnThrThrThrAspValSerGlyPheSerIysGlnArgGlnThrValAlaAlaPhe 60
Db 121 CAACCAACCAACGATGTCAAGTGTTTCTCCAAACAGGACAAACAGTGCGGCGCATTT 180
QY 61 LeuValGluSerGluAsnValIlyLeuLeuValGlyLeuLeuValGlyLeuGluThrVal 80
Db 181 CTGGTGGAGTCTGAAAATGTGAAACTTCTTAAAGATGTCTTAAAGAACTTGAAACAGTA 240
QY 81 ProIlaAsnAsnIlyLeuHisIleValGlnPheAsnGlyProIleLeuGluGluThrIys 100
Db 241 CCGGAAATATATAAATCTCATATGTCCAAATTCAAATGCCCATTTTAAAGAAACAAAA 300
QY 101 GlnIlyLeuGluThrThrGlyValAlaIlyIleLeuAspTyrIleProAspTyrAlaTyrIle 120
Db 301 CAGAAAGCTAGAGACACTCGAGCAAAAGTTCTCGACTACATCCCTGANTTATGCAATATTT 360
QY 121 ValGluTyrGluGlyAspValGlnSerIysValArgSerIleGluHisValGluSerVal 140
Db 361 GTCAGATATGAGGGGAGTGTCAATCAAAAGTCCGCTTCATTTGAAACAGTGGAATCAGTG 420
QY 141 GluProTyrLeuProIlyPheIlyLeuAspProGlnLeuPheThrIlyValIleSerThr 160
Db 421 GAGCCATATCTTGGGAAATACAAATATGATCCCGCTTTTCACAAAAAGGCGCATTCGACG 480
QY 161 LeuValIlyValAlaLeuAspThrIlyGlnAsnAsnValGluValGlnLeuArgGly 180
Db 481 CTGGTAAAGCGGCTTGTGATTCGAAAGCAATATAAAGATGTCATTTAAAGAGCG 540
```

```
QY 181 IleGluGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaIlyPro 200
Db 541 ATCGAGAAATCGCTCAGTACGTGACAGCAATACGCTCATTTTATTAATTCGCAAAAGCCT 600
QY 201 GluTyrIlyValIleMetAsnAspValAlaArgGlyIleValIlyValAlaAspValAlaGlnSer 220
Db 601 GAATATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 221 SerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThr 240
Db 661 AGCTACGCTTTGTATGACAAAGCCAGCATTTGTGCAAGTTCGATTCGATTCGATTCGATTC 720
QY 241 GlyArgAsnAspSerSerMetHisGluValPheArgGlyIlyValIleThrAlaLeuTyrAla 260
Db 721 GGAAGAAACGACAGTTCGATGCAAGAAAGCCTTCGCGGTAAATAACAGCACTATATGCA 780
QY 261 LeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySer 280
Db 781 CTGGGTGAGACGAATTAATGCAATGATACGAACGCTCATGTATCCCATGTGGCAGGTTCCG 840
QY 281 ValLeuGlyAsnGlyValAlaThrAsnIlyGlyMetAlaProGlnAlaAsnLeuValPheGln 300
Db 841 GTATTAGGAATGCGCAGCAATTAAGAAATGCGACCTCAAGCGAATCTGTTTCAA 900
QY 301 SerIleMetAspSerSerGlyIlyLeuGlyIlyLeuProSerAsnLeuGlnThrLeuPhe 320
Db 901 TCCATCATGATGATGACATGTGGGCTTGAGGCTTGACCTTCCATTCGCAACCTTATTC 960
QY 321 SerGlnAlaPheSerAlaGlyValAlaArgIleHisThrAsnSerTyrGlyValAlaValAsn 340
Db 961 AGCCAGCATTCATGACGATGCGCAGAAATTCATACAACTTCGGGGGAGCGCGTGAAT 1020
QY 341 GlyIleTyrThrThrAspSerArgAsnValAspAspTyrValArgIlyAsnAspMetThr 360
Db 1021 GGGGCTACACGACAGATTCAGAAATGTGATGATCATATGAAGAAATGATATACG 1080
QY 361 IleLeuPheAlaAlaGlyAsnGluArgProAsnGlyIlyThrIleSerAlaProGlyThr 380
Db 1081 ATTTCTTTTCGGGCTGGGAAATGAAGGCGCAACGGCGTACCATGATGCACTCGGTACG 1140
QY 381 AlaIlyAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyr 400
Db 1141 GCTTAAACGCGATTAACAGTGGCGCAACGAAACCTGGGTCACAGCTTCGATCTCAT 1200
QY 401 AlaAspAsnIleAsnHisValAlaGlnPheSerSerArgIlyProThrIlyAspGlyArg 420
Db 1201 GCAGATATATTAACAGCTTGGACAGTTCCTTCCCGTGGCCGCAAAAAGATGGCGCA 1260
QY 421 IleIlyProAspValIleAlaProGlyThrTyrIleLeuSerAlaArgSerIleValAla 440
Db 1261 ATCAAGCTCATATGCAAGGCGCAGGACATCAATTTTATCAGGMAATCTTCTCTTGCA 1320
QY 441 ProAspSerSerPheThrAlaAsnHisAspSerIlyTyrAlaTyrMetGlyIlyThrSer 460
Db 1321 CCGATTCCTCTCTCTGCGGCAATCATGACGCAAAATATGCTTATATGGGTGAACGCTCC 1380
QY 461 MetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPheValIlyAsn 480
Db 1381 ATGCAACACCGAATGTGGGGGGAATGTTGCACAGCTCCGTAAGCATTTTGTGAATAAT 1440
QY 481 ArgGlyIleThrProIlyProSerIleLeuLeuValAlaLeuIleAlaGlyAlaAlaAsp 500
Db 1441 AGAGGAATCACTCTTAAGCCTTCCCTTATGAAGCACTTGTGATGAGGCGTGTGAT 1500
QY 501 ValGlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyValArgValThrLeuAspIlySer 520
Db 1501 GTTGATTTGGGTTATCCGAACGAAACCAAGATGGGCGGAGTGACCTCGAATTAATTCG 1560
QY 521 LeuAsnValAlaTyrValAsnGluSerSerAlaLeuSerThrSerGlnValAlaThrTyr 540
Db 1561 TTGAACGTTGCTATGTGGAAGCAATCAGTGCCCTATCAACTGACAAAAGCAACATAT 1620
```

Qy 541 ThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTrpSerAspAlaPro 560
Db 1621 ACCTTACTGCAACGGCGGCAAGCAATTTGAAATCTCCCTGTAAGTCGATGCCCTT 1680
Qy 561 AlaSerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaPro 580
Db 1681 GCAAGACATACGCTCTCTGTAACCCCTGTCATGATTTGGATTGGATTTACAGACCA 1740
Qy 581 AaNGlyThrArgGlyValAlaGlyAsnAspPheSerAlaProPheAspAsnAsnTrpAspGly 600
Db 1741 AACGGAACAGAAATGTGCGAAATGACTTCTCAGCACCAATTGACATATACCTGGATGGC 1800
Qy 601 ArgAsnAsnValAlaGlyAsnValPheIleAsnSerProGlnSerGlyThrTyrrTrileGlu 620
Db 1801 CCCAATTAAGTGAAGAAATGTAATTTATTAATTCGCCCAAGTGAACATATACCTGTTGAG 1860
Qy 621 ValGlnAlaTyrrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1861 GTGCAAGCATATATATGTGCGGTTGGACCAAAACTTCTCGTTGGCAATTTGTAAC 1917

RESULT 2

LOCUS AR368116 1920 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 3 from patent US 6376227.
ACCESSION AR368116
VERSION AR368116.1 GI:34601777
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1920)
AUTHORS Takaiwa, M., Okuda, M., Saeki, K., Kubota, H., Hitomi, J., Kageyama, Y.,
Shikata, S. and Nomura, M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 3 23-APR-2002;
Kao Corporation; Tokyo;
JPX;

FEATURES
source location/Qualifiers
1..1920
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 3,14e-169 Length: 1920
Score: 3267.00 Matches: 639
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-784-870-4 (1-639) x AR368116 (1-1920)

Qy 1 MetArgLysLysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrVal 20
Db 1 ATGAGAAAGAAAGAGAGTGTATTTATCTGTTTATACGCTGAGCAATCTGTCGACTGT 60
Qy 21 AlaLeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIle 40
Db 61 GCATTAAACAATCCCTCGGCTGCTGATGCAAGCACTTTTGACTGGAATTTTAAAGAAATT 120
Qy 41 GlnThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaAlaPhe 60
Db 121 CAAGAAACACCGAAGTCAAGTGTCTTCCAAACAGCAACAAACAGTGGCGCTGCATTT 180
Qy 61 LeuValGlnSerGlyAsnValLysLeuLeuLysGlyLeuLeuLysLysLeuGlnThrVal 80
Db 181 CTGGTGAAGTCTGAAATGTGAATCTTCTTAAGAGATTGCTTAAGAACTTGAACAGTA 240
Qy 81 ProAlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlnGlnThrLys 100
Db 241 CCGCAATTAATTAATCTCAATTTGTCCAAATTCATATGAGCCCATTTTGAAGAAACAAA 300

Qy 101 GlnLysLeuGlnThrThrGlyValAlaLysIleLeuAspTrpIleProAspTrpAlaTyrrIle 120
Db 301 CAGAGCTAAGAGACAACTGAGCAAGAGATTCTCGACTACATCCCTGATTAATGACATATAT 360
Qy 121 ValGlnTyrrGlnGlyAspValGlnSerLysValArgSerIleGlnHisValGlnSerVal 140
Db 361 GTCGAGTAAAGAGGGGGAATTTCAAGTCAAAAGTCCGCTCAATTGAAACCTGGAAATCAGT 420
Qy 141 GluProTyrrLeuProLysTyrrLysIleAspProGlnLeuPheThrLysGlyAlaSerThr 160
Db 421 GAGCAATCTTCCCAATATACAAATATAGTCCCGAGCTTTTCAAAAAGCCCATGACG 480
Qy 161 LeuValLysValAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGlnValGlnLeuAspGly 180
Db 481 CTGGGAAAGCCTTGCGCTTGAATACGAAGCAAAACAAATTAAGATGCAATTAAGAGGC 540
Qy 181 IleGlnGlnIleAlaGlnTyrrValAlaSerAsnAspValHisTyrrIleThrAlaLysPro 200
Db 541 ATCGAGAAATCGCTCAAGTACCTACCAAGCAATGACGCTCATTAATTTACGCAAGCCT 600
Qy 201 GluTyrrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSer 220
Db 601 GAATATTAAGTATGATGATGATGATGCGCCAGAGGATTTGCAAGCGGATGCGACAGAGC 660
Qy 221 SerTyrrGlyLeuTyrrGlyGlnGlnIleValAlaValAlaAspThrGlyLeuAspThr 240
Db 661 AGCTACGGTTGTATGACAAAGCCAGATGTGCGAGTGGCCGAATCTGGAATGGAATACA 720
Qy 241 GlyArgAsnAspSerSerMetHisGlnAlaPheArgGlyLysIleThrAlaLeuTyrrAla 260
Db 721 GGAAGAAACGACAGCTTCATGATGATGATGATGCGCTCCGCGTAAATTAACGACATATGCA 780
Qy 261 LeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySer 280
Db 781 CTGGGCTCGAAGAAATATGCGAATGATGAAACGATCATGATGATGATGATGATGATGAT 840
Qy 281 ValLeuGlyAsnGlyValAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGln 300
Db 841 GTATTAGAAATATGGGCAACGAAATTAAGCAATGACCTCAAGCAACAACTGATTTTCA 900
Qy 301 SerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPhe 320
Db 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Qy 321 SerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyrrGlyAlaAlaValAsn 340
Db 961 AGCGAAGCATTCAGTGAAGTGGCGCAAGATTCATACAACTCGTGGGGGAGCGGTGAT 1020
Qy 341 GlnAlaTyrrThrThrAspSerArgAsnValAspAspTyrrValArgLysAsnAspMetThr 360
Db 1021 GGGGCTTACAGACAGATTCAGAAATGTGATGACTATGTAAAGAAATGATATGACG 1080
Qy 361 IleLeuPheAlaAlaGlyAsnGlyLysProAsnGlyGlyThrIleSerAlaProGlyThr 380
Db 1081 ATTTCTTTCGGGCTGGGAAATGAAGGCCGAAACGGCGGTACCATGATGACCTGTGACG 1140
Qy 381 AlaLysAsnAlaIleThrValGlyAlaThrGlnLeuLeuAspProSerPheGlySerTyrr 400
Db 1141 GCTAAAAACGCAATTAACAGTGGGGCAACCGAAACCTCGGCTCAAGCTTGGTTCCTAT 1200
Qy 401 AlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArg 420
Db 1201 GCAATTAATTAATTAACAGATGACAGATTCCTTCCCTGCGCCCAAAAGATGGGGGA 1260
Qy 421 IleLysProAspValMetAlaProGlyThrTyrrIleLeuSerAlaArgSerSerLeuAla 440
Db 1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Qy 441 ProAspSerSerPheThrAlaAsnHisAspSerTyrrAlaTyrrMetGlyGlyThrSer 460
Db 1321 CCCGATTCCTCTCTTGGGGCAATCATGACAGCAAAATATGCTATATGCGTGAACGTCC 1380
Qy 461 MetAlaThrProIleValAlaGlyAsnValAlaGlnLeuAspGlnHisPheValLysAsn 480

Db	1381	ATGGCAGAACCGGATGGTCCGGGGAATGTGCACGCTCCGGACCATTTGTGAAAAT	1440
Oy	481	ArgGylIeThrProLeuProSerLeuLeuValAlaLeuIleAlaGlyAlaAlaAsp	500
Db	1441	AGAGGAATCATCTCTTAAGCTTCCCTATTGAAAGACGCTTGATGGACGCTGCTCAT	1500
Oy	501	ValGlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrIleValGlyValThrLeuAspLysSer	520
Db	1501	GTTGGATTGGGTTATCCGAACGGAAACCAAGGATGGGGCCGAGTGAACCTGGATTAAATCG	1560
Oy	521	LeuLeuValAlaTyrValAsnGlnSerSerAlaLeuSerThrSerGlnIleValAlaThrTyr	540
Db	1561	TTGAACGTGGCTTAAGTAAACGAATCCAGTGCCTTATCAACTGACCAAAAGCCGACATAT	1620
Oy	541	ThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaPro	560
Db	1621	ACCTTTATCTGCACCGCGCGGACGATTTGAAATCTCCCTGATATGATCGAGTCCCTT	1680
Oy	561	AlaSerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaPro	580
Db	1681	GCAAGCATCACTGCTTCTGTAACTCCGTGCATGATTTGGATTGGTCATTACAGCACCA	1740
Oy	581	AsnGlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGly	600
Db	1741	AACGGAAACAAATATGTCGGGAATACCTTCACGACCAATTTGACAAATTAACGGGATGCG	1800
Oy	601	ArgAsnAsnValGlyAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGlu	620
Db	1801	CGCATATACGTAGAAATGTATTATTATTCGCCCCCAAGTGGAACTATACCTTATAG	1860
Oy	621	ValGlyAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn	639
Db	1861	GTGCAGCATATATATGTCCCGGTTGCACCAAACTTCTCGTTGGCAATTGTGAAC	1917
RESULT 3			
AR562464			
LOCUS	AR562464	1920 bp	DNA linear PAT 08-OCT-2004
DEFINITION	Sequence 3 from patent US 6759228.		
ACCESSION	AR562464		
VERSION	AR562464.1	GI:53976505	
KEYWORDS			
SOURCE			
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1920)		
AUTHORS	Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y., Shikata,S. and Nomura,M.		
TITLE	Alkaline protease		
JOURNAL	Patent: US 6759228-A 3 06-JUL-2004; Kao Corporation; Tokyo, JPY;		
FEATURES			
source	Location/Qualifiers		
	1..1920		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:			
Pred. No.:	3,14e-169	Length:	1920
Score:	3267.00	Matches:	639
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0
US-10-784-870-4 (1-639) x AR562464 (1-1920)			
Oy	1	MetArgLysLysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrVal	20
Db	1	ATGAGAAAGAAAGAGAGGTGTTTATCTGTTTATATACGCTGCAGCAATCTGTGCACTGT	60
Oy	21	AlaLeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIle	40

Db	61	GCATTAAACAATCCCTCGCGTGTATGCGAAGACCTTTGATCTGGATTAAAGSAAT	120
QY	41	GLNTHRTHTTRAspValSerGlyPheSerLysGlnArgGlnThrGlyValAlaAlaPhe	60
Db	121	CMAACACACCGAAGTCAGTGCATGCTTCTCCAAACAGCCACAAACAGTGGCGCTGCATT	180
QY	61	LeuValGlnSerGlnAsnValIleLeuLeuLysGlyLeuLeuLysIleLeuGlnThrVal	80
Db	181	CTGGGTGGAGTCTGAATAATGTGAACCTTCTTAAGCATGCTTAAGAACTTGAAACGTA	240
QY	81	ProIlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlnGlnThrLys	100
Db	241	CCGGCAAAATTAATACTCCATATGTTCCAAATTCAAATGGCCCAATTTTAAAGAAACAAA	300
QY	101	GlnLysLeuGlnThrThrGlnValAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIle	120
Db	301	CAGAGCTAGACACACTGAGCAAAAGATTCTCGACTACATCCCTGATTATATGCAATAT	360
QY	121	ValGlnTyrGlnGlyAspValGlnSerLysValArgSerIleGlnHisValGlnSerVal	140
Db	361	GTGAGTATGAGGGGGATGTTCACTCAAAAGTCCGCTTCATTGAAACACTGGAATCAGTG	420
QY	141	GlnProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThr	160
Db	421	GAGCATATCTGGCCCAAAATACAAATTAATATCCCGACTTTTACAAAGCCGATCGACG	480
QY	161	LeuValLysValAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGlyValGlnLeuArgGly	180
Db	481	CTGGCGAAAGCCTTGCCGCTTGATACGAAGCAAGAACATTAAGAAAGTGCATTAAGAGGC	540
QY	181	IleGlnGlnIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaLysPro	200
Db	541	ATCGAGGAATTCGCTCAGTACGTACCAAGCAATGACGCTCATTAATATTACGGCAAGCCT	600
QY	201	GlnTyrLysValMetAsnAspValAlaArgGlyIleValLysValAspValAlaGlnSer	220
Db	601	GAATTAAGGTAGTAATGATGTGGCCAGAGGTATGTCAAGCGAGTGTGCACAGACC	660
QY	221	SerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThr	240
Db	661	AGCTACGCTTTGTATGACACAGGCCAGATTGTCCAGTGGCCGATACCTGATTGATACA	720
QY	241	GlyArgAsnAspSerSerMetHisGlnAlaPheArgGlyLysIleThrAlaLeuTyrAla	260
Db	721	GGAAGAAACGACAGTTCGATGCAATGAAGCTTCCGCGGTAATAATACACATACATATGCA	780
QY	261	LeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisIleGlyThrHisValAlaGlySer	280
Db	781	CTGGGTCCGAGCAATTAATGCGAATATACGAACGCTCATGTGATCCCATGTGGCAGGTTG	840
QY	281	ValLeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGln	300
Db	841	GTATTAGAAATGGCGCAACGATTAAGGAATGGCACTTCMAAGCGAATCTGGTTTTTCAA	900
QY	301	SerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPhe	320
Db	901	TCCATCATGATAGTAGCAGTGGTGGCTTGGAAGCTTGCTTCACATCTGCAAACTTATTC	960
QY	321	SerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyrGlyAlaAlaValAsn	340
Db	961	AGCCAAAGATTCAAGTCAGAGTCCAGGAATTCATACAACTCTTGGGGGGCAGCGGGAAT	1020
QY	341	GlyAlaTyrThrThrAspSerArgAsnValAspAspTyrValArgGlyAsnAspMetThr	360
Db	1021	GGGGGCTTACGACAGATTCAGAAATGTGATGTGATATGTATGAAAGAAATGATATGACG	1080
QY	361	IleLeuPheAlaAlaGlyAsnGlyLysArgProAsnGlyGlyThrIleSerAlaProGlyThr	380
Db	1081	ATTCTTTTGGGGGCTGGGAATGAAGGCCGAACGGCGGTATCATATGATGACCTGTTACG	1140
QY	381	AlaLysAsnAlaIleThrValGlyAlaThrGlnAsnLeuLysProSerPheGlySerTyr	400

Db 1141 GCTAAACCCGCAATACAGTCCGCGCAACCGAAACCTGCTCAAGCTTCGGTCTAT 1200
Qy 401 AIAAaPaenIlaaahIaValaIaGInPheSerSerArgGlyProThryaaspGlyAaG 420
Db 1201 GGAGAAATATTAACCACTGCGACGTTCTCTCCGCGGCCGCAAAAGATGGCCCA 1260
Qy 421 IleybProaspValaMerlaaProGlyThryrIleLeuSerlaaArgSerSerLeuala 440
Db 1261 ATCAAGCCCTGATGTCAATGCGCCAGGGAACATACCTTTATCAAGAAATCTTCTTGA 1320
Qy 441 ProaPseSerPheThrPaIaahniIaaspSerLyTyraIaTyMeGlyGlyThrSer 460
Db 1321 CCCGATTCCTCTCTGCGCGCAATCAAGCAAAATATGCTTAATATGGGTGAAGTCC 1380
Qy 461 MetAlaThrProIleValaIaGlyAenValaIaGInLeuArgGlyLhiIshValaIyaaen 480
Db 1381 ATGGCAACCCGATGTGTGCGGGAAATGTGCAACAGCTCCGTAAGCATTTTGTGAATAAT 1440
Qy 481 ArgGlyIleThrProIyProSerLeuLeuLySaIaIaLeuIleIaGlyAaIaAaP 500
Db 1441 AGAGAAATCACTCCAGGCTTCCTATTGAAGCAGCTTAATGCGAGTCTGCTGAT 1500
Qy 501 ValGlyLeuGlyTyTProaenGlyAaenGInGlyTTPGlyArgValThreuaapLySer 520
Db 1501 GTTGATATGGGTATATCCGAACGAAACCAAGATGGGCGCAAGTACCTGTGAATAATG 1560
Qy 521 LeuaenValaIaTyTValaIaenGlySerSerlaIaLeuSerThrSerGInLySaIaThryr 540
Db 1561 TTGAACGTGCTTAATGTGAACCAATCCAGTGCCTATCAATAGCAAAACCAAGCATAT 1620
Qy 541 ThrPheThrAlaThraIaGlyLySaProLeuLySaIleSerLeuValThrSerAaP 560
Db 1621 ACCTTTACTGCAACGCGCGGCAAGCCATTGAATACTCCCTGATATGCGAGTCCCT 1680
Qy 561 AlaSerThrThraIaSerValThreuaIaAaenPLeuaapLeuValIleThraIaPro 580
Db 1681 GCAACACACTACTGCTTCTGTAACTCCGTGATGATTTGGATTGGTCACTTCAACACA 1740
Qy 581 AaenGlyThraGlyTyTValaIyAaenAaPheSerlaIaProPheAaPaenAaThraPgly 600
Db 1741 AACGGAACAAATATGTGCGGAATGATCTTCAGACCACTTGAACATTAACGCGATGAC 1800
Qy 601 ArgAaenValaIaAaenValaPheIleAaenSerProGInSerGlyThryrThrIleGIn 620
Db 1801 CGCAATTAAGTAAGAAATGATTAATTAATTCGCCCAAGTGAACATTAACCATTAAG 1860
Qy 621 ValGInaIaTyAaenValaProValaGlyProGInAaenPheSerLeuAlaIleValaen 639
Db 1861 GTGCAAGCATTAATGTGCGGTTGCAACCAAACTTCTGCTTGCAATTGTGAAC 1917

RESULT 4
AR368118 1923 bp DNA linear PAT 12-SBP-2003
LOCUS AR368118 Sequence 7 from patent US 6376227.
DEFINITION AR368118
ACCESSION AR368118
VERSION AR368118.1 GI:34601779
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Seeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 7 23-APR-2002;
Kao Corporation; Tokyo;
JPF;

FEATURES
SOURCE Location/Qualifiers
1..1923 /organism="unknown"
ORIGIN /mol_type="genomic DNA"

Alignment Scores:
Pred. No.: 1,34e-158 Length: 1923
Score: 3072.00 Matches: 536
Percent Similarity: 97.0% Conservative: 23
Best Local Similarity: 93.4% Mismatches: 19
Query Match: 94.0% Indels: 0
Gaps: 0
DB: 6
US-10-784-870-4 (1-639) x AR368118 (1-1923)

Qy 2 ArgLySaLySaValaPheLeuSerValaIleuSerlaaIaAlaIleuSerThraVala 21
Db 7 AAGAAAGAAAGAGTCTTTTATCTGTTTATCAAGTGAAGCATTTGTGCACTGTGG 66
Qy 22 LeuaenAaenProSerlaIaGlyAaPAlaArgThrPheAaPLeuaaPheLyGlyIleGIn 41
Db 67 TTAAGTAATCCATCTGACAGTGTGCAAGAAATTTGATCTGATTCGAAGAAATTCAG 126
Qy 42 ThrThrThraPValaSerGlyPheSerLySaIaArgInThrGlyValaAlaIaPheLeu 61
Db 127 ACAACACTGATGTAAAGGTTTCTCAAGCAGGCGCAAGACTGTGTGCTGCTTTCTG 186
Qy 62 ValGInSerGlyAaenValaIyLeuLeuLySaIyLeuLySaLyLeuGlyThraValaPro 81
Db 187 GTGAAATCTGAATAATGTGAATCTCCCAAAAGGTTTGCAGAAAGAACCTTGAACAGTCCG 246
Qy 82 AlaAaenAaenLyLeuLhiIleValaGInPheAaenGlyProIleLeuGInGlyThryrGIn 101
Db 247 GCAATATATTAATCTCATATATATCAATCAATGACCAATTTTGAAGAAACAAACAG 306
Qy 102 LySaLeuGInThrThrGlyValaIyLeuAaPlyrIleProAaPlyrIaIyIleVala 121
Db 307 CAGCTGAAAGAAACAGGCGCAAGATTCCTCAATCAATACCTGATTAATGCTTAATGTC 366
Qy 122 GluTyGInGlyAaPValaIaGInSerLySaIaArgSerIleGInIaValaIaGInSerValaGIn 141
Db 367 GAGTATGAGGCGATGTTTAAGTCAAGCAACAGACCATTTAGACAGTGTGAATCCGTGAG 426
Qy 142 ProTyLeuProLyTyTyIleAaPProGInLeuAaPheThryrLySaIaAaSerThrye 161
Db 427 CCTTATTTGCGATATACAGATATGATCCCAAGCTTTTCAAAAGGGGCAATCAGAGCTT 486
Qy 162 ValLySaIaLeuAlaIeAaPThryrLySaIyAaenAaenLySaIyAaenAaArgLyIle 181
Db 487 GTAAAGCAAGCGGCTGTGATACAAAGCAAAATTAAGAGGTGCATTAAGAGGCAATC 546
Qy 182 GluGInIleAlaGInTyTValaIaSerAaenAaPValaIshTyrIleThraIaLySaProGIn 201
Db 547 GAACAAATCGCAATTCGCAATTAAGCAATATGTCTATATTAATTCGCGCAAGGCTTAG 606
Qy 202 TyTyLySaIaMeAaenAaPValaIaArgGlyIleValaLySaIaAaPValaIaGInSerSer 221
Db 607 TATAAGGATGATGATGATGTTGCGGTGATTTGCAAAAGGATGTGCTCAAGGAGC 666
Qy 222 TyTyGlyLeuTyTgInGInGInIleValaIaValaIaAaPThryrLySaIyLeuAaPThryr 241
Db 667 TACGGGTGTATGCAAGAGCAAGATCGTACGGGTGCGATACAGGCTTGATACAGGT 726
Qy 242 ArgAaenAaPseSerMeHieGlyAlaPheArgGlyLySaIleThraIaLeuTyraIaLeu 261
Db 727 CGCAATGACAGTGTGATGATGAAGCTTCCGCGGAATAATTAATGCAATTAATGCAATG 786
Qy 262 GlyArgThraAaenAaIaAaenAaPThraAaenGlyAaGlyThrHlIshValaIaGlySerVala 281
Db 787 GAGAGCAAGATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 846
Qy 282 LeuGlyAaenGlyAlaThraAaenLySaIyMeAlaProGInAaAaenValaIaPheGInSer 301
Db 847 TTAGGAACCGCTCACTAATTAAGAAATGCGCTCAGGGAATTAATGCTTCAATCT 906
Qy 302 IleMetAaPseSerSerGlyLyLeuGlyGlyLeuProSerAaenAaenGInThrLeuPheSer 321
Db 907 ATCATGATACCGGTGGGGGACTTGAGAGACTTACCTTCAATCTGCAACCTTAATTCAGC 966

QY 322 GlnAlaPheSerAlaGlyValArgIleHsrThrAsnSerTrpGlyValAlaValAsnGly 341
Db 967 CAAGCATATCAGTGTGGTGGCGAATTCATACAACTCTGGGAGAGCGATGAT33G 1026
QY 342 AlaTyrThrThrAspSerArgGlnValAspAspTyrValArgIleAspAspMetThrIle 361
Db 1027 GCTTACACACAGATTCAGAAATGTGATGCTATGTGGCGAAATGATATGACGATC 1086
QY 362 LeuPheAlaAlaGlyAsnGluArgProAsnGlyIleThrIleSerAlaProGlyThrAla 381
Db 1087 CTTTTCGCTGGCGGAAATGAAGACCGAAGCGGAAACCATCATGACACAGGACAGCT 1146
QY 382 LysAsnAlaIleThrValGlyValAlaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
Db 1147 AAAAAATGCAATTAACAGTGGAGCTACGAGAAACCTCCGCGCAAGCTTGGGCTTATGCG 1206
QY 402 AspAsnIleAsnHleValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIle 421
Db 1207 GACATATACACATGTGGACAGTTCCTTCACTGACCGACCAAGAGATGACGATC 1266
QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
Db 1267 AACCGGATGTCAATGGACCGGGAACGTTCAATCATACAGCAAGATCTTCTTCACACG 1326
QY 442 AspSerSerPheThrAlaAsnHleAspSerLysTyrAlaTyrMetGlyIleThrSerMet 461
Db 1327 GATTCCTTCCTGGGCGCAACCATACAGTAATATGCAATACAGTGGAAACGTCACAG 1386
QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgLysHsrPheValLysAsnArg 481
Db 1387 GCTACACCGATGCTGTGGAAACGTGGACACAGCTTCGAGACATTTGTGAAAAACGA 1446
QY 482 GlyIleThrProLysProSerSerLeuLeuValAlaLeuIleAlaGlyValAlaAspVal 501
Db 1447 GGCATCACACCAACCTTCTCTATTTAAAGCGGACATGATGCGGACGCTACATC 1506
QY 502 GlyLeuGlyTyrTrpAsnGlyValAsnGlnGlyTyrGlyValArgValThrLeuAspLysSerLeu 521
Db 1507 GGCCTTGGCTACCCGACCGTAACCAAGATGGGAGCAAGTGAATGATTAATCCCTCG 1566
QY 522 AsnValAlaTyrValAsnGlyLysSerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
Db 1567 AACGTGCTATGTGAACAGATCCAGTCTCTATTCACACGCAAAAGCGACGTACG 1626
QY 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAla 561
Db 1627 TTATCTGCTACTGCGGCAAGCCTTGAATAATCTCCGTGTATGATGATGCCCTGCG 1686
QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCAACAACCTGCTTCGTAAACGCTTGTCAATGATCGAACCTTGCTATACCCCTCAAT 1746
QY 582 GlyThrArgTyrValGlyValAsnAspPheSerAlaProPheAspAsnAsnTrpAspGlyArg 601
Db 1747 GGCACACAGTATGTGAATATGATCTTACTTCCCATACAAAGATATACCTGGATGGCCG 1806
QY 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
Db 1807 AATAACGTAAAGAAATGATATTATATATGACCAACAAAGCGGAGCTATATCAATTGAAGTA 1866
QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGCTTATTAACGTACCGGTGGACCAAGAACTTCTCGTGGCAATGTGAAT 1920
RESULT 5
AR562466
LOCUS AR562466 1923 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 7 from patent US 6759228.
ACCESSION AR562466
VERSION AR562466.1 GI:53976507
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hicomi,J., Kagayama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6759228-A 7 06-JUL-2004;
Kao Corporation, Tokyo;
JPK;
FEATURES
source Location/Qualifiers
1..1923
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1,34e-158 Length: 1923
Score: 3072.00 Matches: 596
Percent Similarity: 97.0% Conservative: 23
Best Local Similarity: 93.4% Mismatches: 19
Query Match: 94.0% Indels: 0
DB: 6 Gaps: 0
US-10-784-870-4 (1-639) x AR562466 (1-1923)
QY 2 ArgLysLysLysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrValAla 21
Db 7 AAGAGAAAGAGAGGTTTATCTGTTTATCACTGACGCAATTTGTCACTGTTCG 66
QY 22 LeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
Db 67 TTAAATATCATCATCTGACAGTGTGCAAGAAATTTGATCTGATTCGAATTCAAAGAAATTCAG 126
QY 42 ThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyValAlaAlaPheLeu 61
Db 127 ACAACAATGATGTCTAAAGGTTTCTCCAAAGCAGGCGACAGTGTGCTGCTTTCTG 186
QY 62 ValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLysLysLysLysLysLys 81
Db 187 GTGAAATGTGAATATGTGAATCTCCCAAAAGGTTTGCAGAAAGCTTGAAACAGTCCG 246
QY 82 AlaAsnAsnLysLeuHsrIleValGlnPheAsnGlyProIleLeuGlnGluThrLysGln 101
Db 247 GCAATATATTAATCCATATTCATATTCATATTCATATTCATATTCATATTCATATTCAT 306
QY 102 LysLeuGluThrThrGlyValAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
Db 307 CAGCTGAAAAAACAAGGCGCAAAAGATTCCTCACTACATACCTGATTTATGCTTACATGTGC 366
QY 122 GluTyrGluGlyAspValGlnSerLysValArgSerIleGlnHsrValGluSerValGlu 141
Db 367 GAGATATGAGGCGAGATGTATAGTCAACCAACACATATGACACGATGAAATCCGTGAG 426
QY 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThrLeu 161
Db 427 CCTTATTTGGCGGATATACAGATATGATCCACAGCTTTTCAAAAAGGCGCATCGAGCTT 486
QY 162 ValLysValAlaValAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgGlyIle 181
Db 487 GTAAAAAGAGTGGCGCTTGATACAAAGCAAAAAATAAAGAGGTGCATTTAAGGCGATC 546
QY 182 GluGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaLysProGlu 201
Db 547 GAACAAAATCGCAATATGTGCAATTAAGACATATGTCTATATTAACGCAAAAGCTTGAG 606
QY 202 TyrLysValMetAsnAspValAlaArgGlyIleValLysValAspValAlaGlnSerSer 221
Db 607 TATAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
QY 222 TyrGlyLysTyrGlyGlnGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
Db 667 TACGGTTGTATGACAAAGACAGATCGTACGCGTTCGATACAGGCGCTTGAATACAGGT 726

QY 242 ATGAAAPSerSerMetHISGUAAlaPheArgGlyLeuIleThrAlaLeuValLeu 261
 DB 722 CCGCATGACAGTTCATGACGTAAGCTTCGCGGGAATAATCATCATTAATGACATG 786
 QY 262 GATGCTHrAsnAsnAlaAsnAspThrAsnGlyLeuIleValAlaGlySerVal 281
 DB 787 GGAACGACGAATATGATCCAAATGATGCAATGCTCATGATGCTGCTGCTCCGTA 846
 QY 282 LeuGlyAsnGlyAlaThrAsnGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
 DB 847 TTAGGAACCGCTCCCTAATTAAGGAATGCGCCCTCAGCGCAATCTAGTCTTCCAACT 906
 QY 302 IleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSer 321
 DB 907 ATCATGAGATGCGCTGCGGAGACTGAGACATCACTTCGAATCTGCAACCTTAATCAC 966
 QY 322 GlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTrpGlyAlaAlaValAsnGly 341
 DB 967 CAAAGCATACAGTCTGCTGCTCCAGAAATTCATCAAACTCCTGGGAGACACATGAAATGG 1026
 QY 342 AlaIleThrThrAspSerArgAsnValAlaAspArgValArgLeuAsnAspMetThrIle 361
 DB 1027 GCTTACACAAAGATTCAGAAATGATGATGATGCTGCGCAAAATGATATGACGATC 1086
 QY 362 LeuPheAlaAlaGlyAsnGlyLeuArgProAsnGlyGlyIleThrIleSerAlaProGlyThrAla 381
 DB 1087 CTTTTCGCTGCGCGGAATGAAGACCGAAGCGGAAACATCATGTCACAGCGACAGCT 1146
 QY 382 LysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTrpAla 401
 DB 1147 ABAATTCATACAGTCCGACCTACGAAACCTCCGCAAGCTTGGCTTTATGCG 1206
 QY 402 AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIle 421
 DB 1207 GACATATATCAACATGTCGACAGATCTCTTCAACGCGACCGCAAAAGATGACGAGATC 1266
 QY 422 LysProAspValMetAlaProGlyThrTrpIleLeuSerAlaArgSerSerLeuAlaPro 441
 DB 1267 AAACCGAATGATGACCGGGAACGTTCAATCATACGCAAGATCTTCTTGACAGCG 1326
 QY 442 AspSerSerPheTrpAlaAsnHisAspSerLysTrpAlaArgMetGlyIleThrSerMet 461
 DB 1327 GATTCCTCTCTTGGGCGCAACCATGACATTAATGATGATGATGATGATGATGATGATG 1386
 QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlyHisPheValLysAsnArg 481
 DB 1387 GCTACACGATGCTGCTGGAACGTCGACAGCTTCGAGCATTTTGTGAAAAACGA 1446
 QY 482 GlyIleThrProLysProSerLeuLeuValAlaLeuIleAlaGlyAlaAlaAspVal 501
 DB 1447 GGCATCAACCAACCTCTCTATTAATAAGGCGCATGATGTCGCGTCACCTGACATC 1506
 QY 502 GlyLeuGlyTrpProAsnGlyValAsnGlnGlyTrpGlyArgValThrLeuAspLysSerLeu 521
 DB 1507 GGCCTTGGCTACCCGACGTAACGATGAGGAGACGATGACATGATGATGATGATGATG 1566
 QY 522 AsnValAlaArgValAsnGlyLysSerSerAlaLeuSerThrSerGlnLysAlaThrTrpThr 541
 DB 1567 AACGTTGCTTAATGTAACGAGTCCAGTCTCTATCCACGCAAAAGCGACGTACTCG 1626
 QY 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTrpSerAspAlaProAla 561
 DB 1627 TTTACTGCTACGCGCGGACGCTTGAATACTCCCTGCTATGCTGATGCTGATGCTGCTG 1686
 QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
 DB 1687 AGCAACAATGCTTCCGTAAAGCTTGTCAATGATCTGACCTTGTCTCAATCCCTCCAAAT 1746
 QY 582 GlyThrArgTrpValGlyAsnAspPheSerAlaProPheAspAsnAsnTrpAspGlyArg 601
 DB 1747 GGCACACAGTATGTAAGAAATGATCTTACTTCCCATCAATGAATGATGATGATGATG 1806
 QY 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTrpThrIleGluVal 621

DB 1807 AATAACGTAAGAAATGATTTATTAATGACCAACAAAGCGGACGATACATTAAGTA 1866
 QY 622 GlnAlaArgAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
 DB 1867 CAGGCTTAATACGTAACCGGTTGACCAAGAACTTCTGTTGGCAATTTGTGAAT 1920
 RESULT 6
 AB051423
 LOCUS
 DEFINITION
 AB051423
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLES
 JOURNAL
 REFERENCE
 AUTHORS
 TITLES
 JOURNAL
 COMMENT
 FEATURES
 source
 location/Qualifiers
 1..1923
 /organism="Bacillus sp. KSM-KP43"
 /mol_type="genomic DNA"
 /strain="KP43"
 /db_xref="taxon:109322"
 1..1923
 /gene="PROF"
 1..1923
 /gene="PROF"
 /codon_start=1
 /transl_table=11
 /product="protease"
 /protein_id="BAB5674.2"
 /db_xref="GI:20521155"
 /translation="MRKKKKYFLSVLSAAALSTVALSNPSGARNFDLDFKGIQTT
 TDAGFSKGGTGAAPLVESENVGLPGGLKSTVPANNKLAHIFONGPILBETKO
 QLBKGAKLIDYIDVAYIVRYBGQVNSATSTIHESVEPLPIVRIIDPOLFTGAS
 ELVRAVALDTQKKKEVQLRGIHQIAQPAISNDVLYITAKREYKRMNDVAKIVADV
 AOSSTGLTGQGIIVAVADTGLDTGKNDSSHEAFRGKITALVALGRTNNANDTNGHT
 HVASVILNGSTINGMAPOANLVFQSIINDSGGLGGLPSNLTQLFSGQASAGARLHTN
 SWGAIVNGAYITDSRNVDDYVRKNDMTILFAAGNEPNGCTISAGTAKNAITVATN
 NLKPSFGYADNINHVAFSSRGPTKRIEDVVAPEGTFILSASSLAPDSFWAN
 DSKAYVNGSTQKATPIVAGVNAOLREHVRKRGITPKSLKALIIAQAADIGCYPN
 GNQGRVTLTKSLNAVYVNESSISLSTQKATYTSFTATAGKPLKISLWSDAPASTTA
 SVTLVNDLIDLVTPNGTQYVGNDFTSYNDNWDGRNNVENVFVINAPOSGYTILEVQA
 YNVVPGQTFSLAIYN"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.85e-158 Length: 1923
 Score: 3066.00 Matches: 595
 Percent Similarity: 96.94 Conserved: 23
 Best Local Similarity: 93.34 Mismatches: 20
 Query Match: 93.84 Indels: 0
 DB: 1 Gaps: 0
 US-10-784-870-4 (1-639) x AB051423 (1-1923)
 QY 2 ArgLysLysLysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrValAla 21
 DB 7 AAGAAAGAAAAGTGTTTTATCTGTTTATACAGTGCAGGATTTTGTGCACTGTTGCG 66
 QY 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41


```
/cranal table=11
/protein_id="CA61612.1"
/db_xref="GI:54304587"
/translation="MRKKKVFSLVLSAAAILSTVALSNPSAGANFOLDPFGIGIT
TDAKFSKQGTGAFAFLVESBNVKLPKLOKLETPVANNKHLIQFNGPLLEETKO
QLEKFGAKLIDYIPDAVIVYEGDVKASTSTLEHVSVEPLPIRIRIDPOLTEKAS
ELVKVVALDTKQNKKEVOLRGILOIAQPSINDVLYTKPKRYKWNDAIRGIVKADY
AGSYGLVGCGQIVVADPTGIDPTGRDSSMHEAPRGKITRLVLAQRTNANNNGHGT
HYAGSYLNGSTRKGMARPOANLVPQSIHSGGGLGGLPSNLQTLFSQATSAQARLHTN
SWGAANVAGYTTDSRNVDYVRKNDMTLLPAAHNEBPNGGTLIAPGTAKNALTVGATE
NLPSPGSAVDININHAQPSRSGPTDGGIKPDVAPGTFPIISARSISLAPDSFMANH
DSKVAHMGSTMAATPIVAGNVAQLREHFYKNGRIKTPKPLKALALAGADJGLGPN
GQNGMRVTLDKSLNVAYNSSSLSTSKATYSPATKPKPKLKTSLVNSDAPASTA
STLYVNDLDTVTTPNGTQYVANDPTSPINDMDGNANVENPYNAPQSGTTTIEVQA
YVAVGPQTFSLAIYV"
1.618
619
```

Alignment Scores:

Prod. No.:	2.85e-158	Length:	1923
Score:	3066.00	Matches:	595
Percent Similarity:	96.9%	Conservative:	23
Best Local Similarity:	93.3%	Mismatches:	20
Query Match:	93.8%	Indels:	0
DB:	6	Gaps:	0

US-10-784-870-4 (1-639) x CQ888285 (1-1923)

```
QY 2 ARGVLYLYLYVALPHEUSERVALLEUSERALALALALEUSERTHRVALA 21
DB 7 AAGAGAAAGAGGTGTTTATCTGTTTATCAGCTGCAGCAATTGTGACGTGG 66
QY 22 LEUAAANPPOSERALAGLYAPRALAARGTHRPHEASPLEUAPHELYSGYLEG 41
DB 67 TTAAGTAATCCATCTGCAGGTGGTGCAGAGAAATTTTGATCTGCAATTTCAA 126
QY 42 THRTHRTHAPRYALSERGLYPHESELYSGINARGINTHRYALALAPHELEU 61
DB 127 ACAACAACTGATCTAAAGGTTTCTCCAAAGCGGCGACACTGGTCTGCTTTTC 186
QY 62 VALGUSERGLYANVALYLYLEULEUYLYLEULEYLYLYLEUGINTHRYALPRO 81
DB 187 GCGAATCTGAATATGTGAATCTCCAAAGGTTTCAGAAAGAGTTGAACAGTCCG 246
QY 82 ALAABHLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 101
DB 247 GCAATTAATACTCCATTAATCCAAATTCATGACCAATTTTAAGAAACAAACG 306
QY 102 LYLEUGINTHRYALYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 121
DB 307 CAGCTGGAAAAACAGGCGCAAGATTCCTGACTACATCTGATTAATGCTTAC 366
QY 122 GIUYTGLUGLYAPRYALINSELYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 141
DB 367 GAGTATGAGGGGAGTGTAAAGTACGACAAAGCACTGAGGACCTGAGGAG 426
QY 142 PROTYLEUPROLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 161
DB 427 CCTATTGTCGCAATACAGATAGATCCCAAGCTTTTCAAAAGGCGCATCAGCT 486
QY 162 VALLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 181
DB 487 GTAAAAAGCGTGGCGCTGATACAAAGCAAAAAAATAAGAGTGCATTAAGA 546
QY 182 GIUGLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 201
DB 547 GAACAAATGCAAAATTCGAATACGAATGATGCTATATATACGCGAAAGCTG 606
QY 202 TYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 221
DB 607 TTTAAGGTATGATGATGCTGGCGCTGGAATTTGTCAAAAGCGGATGCTG 666
```

```
QY 222 TYRGLYLEUTYRGLYGLNGLYGLNILEVALIYALALASPRTGTYLEUAPRTG 241
DB 667 TACGGGTGTATGACAAAGACAGATCGTAGCGGTGCCGATACAGGCGCTTGAT 726
QY 242 ARGAAANPSESESETHNLEGLYALAPHEARGLYLYLYLYLYLYLYLYLYLYLY 261
DB 727 CCGAATGACAGTTGCAATGATGAGCTTCCGCGGGAATAATACATTAATGATG 786
QY 262 GLYATGTHAPRYALAPRYALAPRYALAPRYALAPRYALAPRYALAPRYAL 281
DB 787 GACCGGCAAAATATGCCAATGATACGAATGATGATGATGATGATGATGATG 846
QY 282 LEUGLYANGLYALYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 301
DB 847 TTAGGAAACGCTCCCAATTAAGGAAATGCGCTCAGCGCAATCTATCTTCA 906
QY 302 ILEMEAPSESESESESESESESESESESESESESESESESESESESESE 321
DB 907 ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
QY 322 GINALAPHESESESESESESESESESESESESESESESESESESESESE 341
DB 967 CAGCAATACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
QY 342 ALATYRTHRTHAPRYALAPRYALAPRYALAPRYALAPRYALAPRYALAPRY 361
DB 1027 GCTTACACAAAGATTCAGAAATGATGATGATGATGATGATGATGATGAT 1086
QY 362 LEUPHEALALAGLYANGLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 381
DB 1087 CTTTTCGCTCCGGAATGATGATGATGATGATGATGATGATGATGATGAT 1146
QY 382 LYAANALALYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 401
DB 1147 AAAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1206
QY 402 AAPAANLEAHNILEVALIYALINPHESESESESESESESESESESESESE 421
DB 1207 GACAAATATCAACATGATGATGATGATGATGATGATGATGATGATGAT 1266
QY 422 LYSPROAPRYALAPRYALAPRYALAPRYALAPRYALAPRYALAPRYALAP 441
DB 1267 AAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
QY 442 AAPSESESESESESESESESESESESESESESESESESESESESESESE 461
DB 1327 GATTCCTCTCTGCGGCAACATGATGATGATGATGATGATGATGATGAT 1386
QY 462 ALATHTPROILEVALIYALYANVALIYALINLEUARGGLYLYLYLYLYLY 481
DB 1387 GCTAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
QY 482 GIYILETHPROLYPROSESESESESESESESESESESESESESESESESE 501
DB 1447 GGCATACACCAAAAGCTTCTTATTAAGCGGCACTGATGCTGCTGCTGCT 1506
QY 502 GIYLEUGLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 521
DB 1507 GGCCTTGGCTACCGGAGGATGATGATGATGATGATGATGATGATGATGAT 1566
QY 522 AANVALIATYRVALANGLYSESESESESESESESESESESESESESESESE 541
DB 1567 AACGTTGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1626
QY 542 PHEATHALYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 561
DB 1627 TTTACTGCTACTGCGGCAAGCTTGTAAATCTCCCTGATGATGATGATG 1686
QY 562 SEETHRTHALYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLYLY 581
DB 1687 AACCAACAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1746
QY 582 GLYTHARGTYRVALIYANPHESESESESESESESESESESESESESESE 601
```

|||||.....
Db 1747 GGCACACAGTATGTAGAAATGACTTACTTCCCATACAAAGATACCGGATGCGCC 1806
Oy 602 AaenavaiGluAenValPheIleuSerProGlnSerGlyThrThrIleGluVal 621
Db 1807 AATAAGTAAAGAAATGATATTTATATATGACCAAGAGCGGAGCGATATGCAATTGAGGTA 1866
Oy 622 GlnAaIeThrAenValProValGlyProGlnAenPheSerIleuAlaIleValAen 639
Db 1867 CAGGCTTAAACGTAACCGGTGTGACCAAGACCTTCTCGTTGCAATGTGTGAAT 1920
RESULT 8
CO889309 1923 bp DNA linear PAT 19-OCT-2004
LOCUS Sequence 2 from Patent EP1466962.
ACCESSION CO889309
VERSION CO889309.1 GI:54305176
KEYWORDS
SOURCE Bacillus sp. KSM-KP43
ORGANISM Bacillus; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Okuda, M., Sato, T., Takimura, Y., Sumitomo, N. and Kobayashi, T.
TITLE Alkaline protease
JOURNAL Patent: JP 1466962-A 2 13-OCT-2004;
Kao Corporation (JP)
FEATURES
source location/Qualifiers
1..1923
/organism="Bacillus sp. KSM-KP43"
/mol_type="unassigned DNA"
/db_xref="taxon:109322"
1..1923
/note="unassigned protein product"
/codon_start=1
/transl_table=1
/protein_id="CAH61686.1"
/db_xref="GI:54305177"
translation="MRKKKVKVLSLSAAALISTVALNSPSAGARNPDLPFKIQTT
TDAPKPSKQGTGAAAPVBSSENVKLPKGLQKLETVANPKKHLIQRPGLBETKO
QLEKGAKLDTIPDAIVYVEBVDKATSTIEHVESVEPLPIYRDPQLEFKGAS
ELVAVAVLDTRQKNEVQLRIGIQAPALISNDVLYITAKPEYKVMNDAGIVADY
AOSVYGLGQGVAVADTGLDGTGNDSSMEAPFKGKITALYALGRTNNADNGHGT
HVAGSVLNGSTNKMAPOANLVFQSIIDSGGLGGLPSNLQTLPSQAYSAGARLHTN
SMGAVNAGATTTDSRNVDDYVRKNDMTLLPAGNEGPRGGTISARGTAKNLTVGATE
NLBPSFGSYADNINHVAPSSRGPTRKDRIDEDVMAPEFTLSASSLAPOSSFRAN
DSKATVNGTSMATIVAGNVAQLEKHTVKNRGITPKPSLEKALIAAGADIGLEFN
GNQGWGRVTLDSLVNAVAVNESSLSIQKATYSTATAGKELKLSLVMSDAPASTTA
SVTLVNDLIDLVITAEVNGTQYVGNDFSPYNDMDGRNVENVFINAPOSQTYTTEVQA
NVNPGPQTSPSALIVN"
1..618
sig_peptide
619..1920
mat_peptide
/product="unassigned"
ORIGIN
Alignment Scores:
Pred. No.: 2,856-158 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.94 Conservative: 23
Best Local Similarity: 93.34 Mismatches: 20
Query Match: 93.84 Indels: 0
Gaps: 0
US-10-784-870-4 (1-639) x CO889309 (1-1923)
Oy 2 ArgIyBlyBlyValPheIleuSerValIleuSerAlaIleuSerThrValAa 21
Db 7 AAGAGAAAGAGTGTATTATCTGTTTATCAGCTGCAGCATTTTGCAGCTGTGCG 66
Oy 22 LeuAenAenProSerAlaGlyAaPaIaargThrPheAapIeuaAaPheIyIleGln 41
Db 67 TTAAGTAATTCATTCGAGGTGTGCAAGGAATTTTGATGATTTCCAAAGAAATTCG 126
Oy 42 ThrThrAaPvaIseRcIyPheSerIyGlnArgIInThrGlyAlaIleuAaPheIu 61

|||||.....
Db 127 ACAACACTGATGCTTAAAGCTTTCCTCAAGAGGGGCGAGACTGGTGGCTGCTTTCTG 186
Oy 62 ValGluSerGluAenValIleuIleuIyGlyIleuIyBlyLeuGluThrValPro 81
Db 187 GTGGAATCTGAAGATGTGAATCTCCCAAGAGTTTGCAGAGAAAGCTTGAAACAGTCCCG 246
Oy 82 AlaAenAenIyBlyIleuAlaIleuAlaIleuAenIyProIleuIyGluIleuThrIyGln 101
Db 247 GCAATTAATTAATCTCAATTAATCAATTCATTAAGGACCAATTTTAAAGAAACAAACAG 306
Oy 102 LysLeuGluThrThrIleuAlaIleuIleuAaPThrIleuProAaPThrAlaIleuAla 121
Db 307 CAGCTGGAAGAAACAGGGGCAAGATTTCTCAATCAATCAATCAATTAATGCTTACATTTTC 366
Oy 122 GluIyGluIyAaPvaIleuIyGlnSerIyValIyGlnSerIleuIleuAlaIleuIyGln 141
Db 367 GAGTATGAGGCGCATGTTAAAGTACAGCAACAGCACCATTTAGACGCTGGAATCCGTGAG 426
Oy 142 ProIyIleuProIyIyIleuIleuAaPProGlnIleuPheThrIyGlyAlaSerThrIleu 161
Db 427 CTTATTTGCGATATACAGATATGATCCAGCTTTTCAAAAGGGGCAATCAGAGCTT 486
Oy 162 ValIyAaIleuAlaIleuAaPThrIyGlnAaAenIyGluValGlnIleuAaGlyIle 181
Db 487 GTAAAGCAAGTGGCGCTTGATACAAAGCAAGAAATTAAGAGTGCATTTAAAGAGCATC 546
Oy 182 GluIuIleuIleuIyGlnIyGlnIleuAlaIleuAaPThrIyIleuIleuAaPProGlu 201
Db 547 GAACAAATCGCAATTCGCAATTAAGCATATGATGCTATTAATTAATTCGCAAGCTGAG 606
Oy 202 TyrIyAaValIleuAaAaPvaIleuAaGlyIleuValIleuAaPvaIleuIleuSer 221
Db 607 TATTAAGGATGAATGATGTTGCGGTGAATTTGCAAGGAGAGTGGCTCAGAGCAGC 666
Oy 222 TyrGlyIleuIyGlyGlnIyGlnIleuAlaIleuAaPThrIyIleuAaPThrGly 241
Db 667 TACGGTGTGATGACAAAGACAGATGCTACGGTGGCTGCAATCAGGGCTTGATACAGGT 726
Oy 242 ArgAaAaPSerSerMetHisGluAlaPheAaGlyIyIleuAlaIleuIyIleuAa 261
Db 727 CGCAATGACATGCTGATCAGAAAGACCTTCCGCGGAAATTAATGCTTATATGCAATG 786
Oy 262 GlyIyGThrAaAaAaAaAaAaPThrAaGlnIyIleuIyIleuIleuIleuAlaIleuIleu 281
Db 787 GACGAGCAATTAATGCAATGATACATGATGATGATGATGATGATGATGATGATGATGAT 846
Oy 282 LeuGlyAaGlyAlaIleuAaAaPThrAaGlnIyIleuAaPThrIleuAaPThrIleu 301
Db 847 TTAGGAAACGGCTCCATTAATTAAGAAATGCGGCTCAGGCGAATCTTCCAAATCT 906
Oy 302 IleuAaPSerSerGlyIyLeuGlyIyIleuProSerAaAaPThrIleuPheSer 321
Db 907 ATCATGATATGCGGTGGGAGCTTGCAGAGACTTCTGAAATCTGCAACCTTATTCAGC 966
Oy 322 GlnAaPheSerAlaGlyAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 341
Db 967 CAGAGATACATGCTGCGGCAAGATTAATCAAACTCTGGGAGCGAGCTGAATGG 1026
Oy 342 AlaIyThrThrAaPSerAaAaPvaIleuAaPThrIleuAaIleuAaPThrIleuAa 361
Db 1027 GCTTACCAACAGATTCAGAAATGTGTGATCATATGCGCAAAAGATATGACATC 1086
Oy 362 LeuPheAlaIleuAaIleuAaPThrIleuAaPThrIleuAaPThrIleuAaPThrIleu 381
Db 1087 CTTTTCGCTGCGGGAATGAAGGACCGAAACCGGGAACCAATCAATGACAGGACAGCT 1146
Oy 382 LysAaAaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 401
Db 1147 AAAAATGATTAACAGTGGAGCTTACGGAACCTCGGCAAGCTTGGGATTTATGCG 1206
Oy 402 AaAaAaIleuAaIleuAaIleuAaPheSerAaPThrIleuAaPThrIleuAaPThrIleu 421

Db 1207 GACAAATATCAACCATGTGGACAGCTTCTCTTCACGTGACCGCAAGATGACCGATTC 1266
QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerLysArgSerLeuAlaPro 441
Db 1267 AAAACCGAATGTCATGGACCGGAGAGTTTCATATACAGCAAGATCTTCTTGGACCG 1326
QY 442 AspSerSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMetGlyGlyThrSerMet 461
Db 1327 GATTCCTCTTCTGGCGCAACCAATGCAATTAATGCAATGAGGTGGAACTGTCATG 1386
QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlyHisPheValIlyAsnArg 481
Db 1387 GCTAACCCGATCGTGTCTGAAAACGTGGACAGCTTCGTAGCACTTTGTGAAAACACAG 1446
QY 482 GlyIleThrProLysProSerLeuLeuValAlaIleuIleAlaGlyAlaAlaAspVal 501
Db 1447 GGCATCACACCAAGCTTCTCTATTAAGCGGCACTGATGGCCGTGACGTGACATC 1506
QY 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeu 521
Db 1507 GGCCTTGGCTACCCGACAGGTAAACAGATGGGACAGTGCATGATTAATCCCTG 1566
QY 522 AsnValAlaTyrValAsnGlySerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
Db 1567 AACGTTGCCTATGTGAACAGATCCAGTTCTATCCACAGCCAAAAGCAGTACCTG 1626
QY 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValIleThrSerAlaProAla 561
Db 1627 TTTACTGCTACACGCGGCAAGCTTGAAGATCTCCCTGTATGCTGATGCCCTGCG 1686
QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCAACATCGCTCCCTGATACGCTTGTCAATGATCTGGACCTTGTCACTCCGCTCCAAAT 1746
QY 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTriAspGlyArg 601
Db 1747 GGCACACAGTATGTAGAAATAGACTTACTTCGCCATACATGATTAACCTGGATGGCCG 1806
QY 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
Db 1807 AATAACGTAAGAAATGATATTATTAATGACACCAAGCGGACGTATCAATTGAGCTA 1866
QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGCTTATACGTAACCGGTGGACCAAGACCTTCTCGTGGCAATGTGAAT 1920

RESULT 9
AR368117 1923 bp DNA linear PART 12-SRP-2003
LOCUS AR368117
DEFINITION Sequence 5 from patent US 6376227.
ACCESSION AR368117
VERSION AR368117.1 GI:34601778
KEYWORDS
SOURCE .
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeaki,K., Kubota,H., Hitomi,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 5 23-APR-2002;
Kao Corporation; Tokyo;
JPK;
FEATURES
source Location/Qualifiers
1..1923
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 2,85e-158 Length: 1923
Pred. No.: 3066.00 Matches: 595
Score: 3066.00
Percent Similarity: 96.9% Conservative: 23

Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 6 Gaps: 0
US-10-784-870-4 (1-639) x AR368117 (1-1923)

QY 2 ArgLysLysLysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrValAla 21
Db 7 AAGAAAGAAAAGGTGTTTATCTGTTTATACGCTGCAGCCATTTTGTGACTGTGGCG 66
QY 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
Db 67 TTAAGTAATCCATCTGCAGGTGTGTCAAGAAATTTTGATCTGGAATTCAGAAAGAAATTCAG 126
QY 42 ThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaIlePheLeu 61
Db 127 ACNACACATGATGCTTAAGTTTCTCCAAAGCGGGCAGACGTGCTGCTCTTTTCG 186
QY 62 ValGlySerGluAsnValIlyLeuLeuLysGlyLeuLeuValIleLeuGlnThrValPro 81
Db 187 GTGAATCTGAATAATGTGAACCTCCAAAGGTTTGCAAGAGAGCTTGAAACAGTCCG 246
QY 82 AlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlnGlnThrLysGln 101
Db 247 GCAATTAATTAATCTCATATTAATCCAAATTGATGACCAATTTTGAAGAAACAAACAG 306
QY 102 LysLeuGlnThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
Db 307 CAGCTGAAAAAAGCGGGCAAGATCTCGACTCATACATGATTAATGCTTAACATTCG 366
QY 122 GlyTyrGlnGlyAspValGlnSerSerValAsnSerIleGlnHisValGlnSerValGln 141
Db 367 GAGTATGAGGGGAGATGTTAAGTACAGCAACAGACCATTTGACAGTGAATCCGCGAG 426
QY 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThrLeu 161
Db 427 CCTTATTTGGCATATACAGATATAGATCCCAAGCTTTTCACAAAAGGGGCATCAGAGCTT 486
QY 162 ValIlyValAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGlnValGlnLeuArgGlyIle 181
Db 487 GTAAAAAGACAGTGGCCCTTGATCAAGCAAGAAATTAAGAGGTCAATTAAGAGCATC 546
QY 182 GlnGlnIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaLysProGln 201
Db 547 GAACAATCGCAATATGCAATTAAGCAATGATGTGCTATTAATTAACGCAAGGCTGGAG 606
QY 202 TyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSer 221
Db 607 TTTAAGTGATGATGATGATCTTCCGCTGGAAATTCGAAAGCGGAGTGTGCTCAGAGCAGC 666
QY 222 TyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
Db 667 TACGGGTGTATGACCAAGACAGATCGTAGCGGTGCGATACAGGGCTTATACAGGT 726
QY 242 ArgAsnAspSerSerMetHisGlnAlaPheAspGlyLysIleThrAlaLeuTyrAlaLeu 261
Db 727 CGCAATGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlnHisIleGlyThrHisValAlaGlySerVal 281
Db 787 GAGCGGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 846
QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
Db 847 TTAGAAACCGGCTCACTTAATTAAGAAATGCGGCTCAGCGGAATCTAGTCTTCAATCT 906
QY 302 IleMetAspSerSerGlyLysGlyLeuGlyLysLeuProSerAsnLeuGlnThrLeuPheSer 321
Db 907 ATCATGATAGAGGAGTGGGAGACTTGGAGAGACTTCTGAAATCTGCAAACTTATATTCAGC 966
QY 322 GlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyrGlyAlaAlaValAsnGly 341
Db 967 CAGCATACAGTGTGCTGCTGCGCAAGATTCATCAAACTCTGGGAGAGCAGTGAATGGG 1026

Qy 342 AAlaTyrThrThrAspSerArgAsnValAspAspTyrValArgValAsnAspMetThrIle 361
Db 1027 GCTTACACAAACGATTCAGAAATGTGATGCTATGTGGCGGAAAAAGATATGACGATC 1086
Qy 362 LeuPheAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAla 381
Db 1087 CTTTTCGCTGCGGAAATGAAAGACCGAACCGGAAACATCATGTCACAGGCACT 1146
Qy 382 LysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
Db 1147 AAAAATGCAATACAGTCGAGCTACGGAAACCTCCGCCAAAGCTTTGGGCTTATGCG 1206
Qy 402 AsnAsnIleAsnHleValAlaGlnPheSerSerAlaGlyProThrLysAspGlyValGlyIle 421
Db 1207 GACATATCAACCAAGTGGACAGTTCTTCTTCACTGACCGCAAAAGATGGACGGATC 1266
Qy 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
Db 1267 AAACCGATGTCATGCGACCGGAAACGTTCACTATCATGCAAGATCTTCTTTCACCG 1326
Qy 442 AspSerSerPheTyrPalaAsnHleAspSerLysTyrAlaTyrMetGlyGlyThrSerMet 461
Db 1327 GATTCCTCTCTGCGGCAACCATGATGATTAATGATCATGCGTGGAAACGTCGATG 1386
Qy 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHlePheValLysAsnArg 481
Db 1387 GCTACACCGATCGTGTGGAACCGTGGACAGCTTCCTGAGCATTTGTGAAAAACGA 1446
Qy 482 GlyIleThrProLysProSerLeuLeuLysAlaIleValAlaGlyAlaAlaAspVal 501
Db 1447 GGCAATCAACCAAGCCTTCTCTATTAAGGGGCACTGATTTGCGGTGCACTGACATC 1506
Qy 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeu 521
Db 1507 GGCCTTGGCTACCGGAACGATACCAAGGATGGGACAGTGAATGGATTAATCCCTG 1566
Qy 522 AsnValAlaTyrValAsnGluSerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
Db 1567 AACGTGGCTTATGGAACGATGCAAGTCTCTATCCACCAAGCAAAAGCGATACCTG 1626
Qy 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAla 561
Db 1627 TTTACTGCTACTGCGGCAAGCTTTGAAATCTCCCTGGTATGCTGATGCCCTGCG 1686
Qy 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCACTACTGCTTCGTAACGCTGTGATGATCTGACACTTGCATTACCGCTCAAT 1746
Qy 582 GlyThrArgTyrValGlyAsnAspPheSerSerAlaProPheAspAsnAsnTyrAspGlyArg 601
Db 1747 GGCACACAGTATGTAAGAAATGACTTCTTCGTCATCAAGATACCTGGGATGGCGCG 1806
Qy 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
Db 1807 AATTAACGTAAAGAAATGATTTTATATGACACCAAAAGCGGACGTATTAACAATTGAGTA 1866
Qy 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGCTTATTAACGTACCGGTGGACCAAGACCTTCTCGTTGGCAATTGTGAT 1920

RESULT 10
ARS62465 LOCUS ARS62465 1923 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 5 from patent US 6759228.
ACCESSION ARS62465
VERSION ARS62465.1 GI:53976506
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Takeiwa,M., Okude,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,

TITLE Shikara,S. and Nomura,M.
JOURNAL Alkaline protease
Patent: US 6759228-A 5 06-JUL-2004;
Kao Corporation, Tokyo;
JFX;
FEATURES location/Qualifiers
source 1..1923
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 2,85e-158 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatch: 20
Query Match: 93.8% Indels: 0
DB: 6 Gaps: 0
US-10-784-870-4 (1-639) x ARS62465 (1-1923)
Qy 2 ArgLysIleLysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrValAla 21
Db 7 AAGAAAGAAAAGGTGTTTATCTGTTTATCACTGACGCGATTGTGCACTGTGCG 66
Qy 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
Db 67 TTAAGTATTCATCTGCGAGGTGGACGAAGATTTGATCTGATTTCAAGAAATTCAG 126
Qy 42 ThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaAlaPheLeu 61
Db 127 ACAACACTGATGCTAAAGTTTCTCCAGCAGGGGCACTGCTGCTTCTG 186
Qy 62 ValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLysGlyIleThrValPro 81
Db 187 GTGGAATCTGAAATGTGAATCTCCAAAGGTTTGCAGAAAGCTTGAACAGTCCG 246
Qy 82 AlaAsnAsnLysLeuHleIleValGlnPheAsnGlyProIleLeuGluGluThrLysGln 101
Db 247 GCAATATATTAACCTCCATATTATCAATTCATGACCAATTTTAAAGAAACAAACAG 306
Qy 102 LysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
Db 307 CAGCTGAAAAAACAAGGGGCAAAATCTGCATCACTCACTGATTAATGCTTACATGTC 366
Qy 122 GluTyrGluGlyAspValGlnSerLysValArgSerIleGluHleValGluSerValGlu 141
Db 367 GAGTATGAGGGCGATGTTAAAGTCAGCAACAGCACATTCAGACGTCGATCCGTGAG 426
Qy 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThrLeu 161
Db 427 CTTATTTTGGCAATATCAAGATATGATCCAGCTTTTCAAAAGGGGCACTCAGAGCTT 486
Qy 162 ValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgGlyIle 181
Db 487 GTAAGAGAGTGGCGCTTGAATCAACAGCAAAATTAAGAGTGCATTAAGAGGATC 546
Qy 182 GluGluIleAlaGlnTyrValAlaSerAsnAspValHleTyrIleThrAlaLysProGlu 201
Db 547 GAAACAAATCCGACATTCGCAATTAAGCATGATGTGCTATATTAACGCAAAACCTGAG 606
Qy 202 TyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSer 221
Db 607 TATTAAGTGATGAATGATGTGGCGGTGAATTTGTCAAAAGCGATGTGCTCAAGACAGC 666
Qy 222 TyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
Db 667 TACGGTTGTATGACAAAGACAGATCGTACCGGTGCGGATCAAGGGCTTGAATACAGCT 726
Qy 242 ArgAsnAspSerSerMetHleGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeu 261
Db 727 CGCATGACAGTTGATGATGATGAAGCTTCCGCGGAAAAATTAATGCACTATATATGCAATTG 786


```
Db 187 GTGGAAATGTGAAAAGTGAACCTCCCAAAAGTTTGCAGAGAAACCTTGAACCTCCG 246
Qy 82 ALaasnaenlybLeuHnIleValGlnpheannglyProIleLeuGluGluThrlySerGln 101
Db 247 GCAATATATTAACCTCCATATTTGTCATTCATGAGCAATTTTGAAGAAACAAAGCAG 306
Qy 102 LysLeuGluThrThrGlyValAluylIleLeuSerTyrlIleProSerTyrlAlaThrIleVal 121
Db 307 CAGCTAGAAAACAGGGGCAAAAGATTCCTGACCTACATACCTGATTAAGCTTACATTCG 366
Qy 122 GluTyrgluGluYAspValGlnserTybValArgserIleGluHnIleValGlnSerValGlu 141
Db 367 GAGTATGAGGGGCGAGCTTAAGTCACAGCAACAGCACCATTTAGAGAGCTGGAAATCCCTAAG 426
Qy 142 ProTyLeuProLyTyTylybIleAspProGlnLeuPheThrlyGlyValAserThrleu 161
Db 427 CCTTATTTGGCCGATATACAGATAGATCCCGAGCTTTTCACAAAAGGGGCAATCAGAGCTT 486
Qy 162 VallybAlaLeuAlaLeuSerThrlybGlnaenblybGluValGlnLeuHnrglyIle 181
Db 487 GTAAAAGAGTGGCCCTTGATACAAATCAAAAATAAAGAGGTCCAAATTAAAGGCAATC 546
Qy 182 GluGluIleAlaGlnTyTyValAlaIaserAsnaerValHnIleTyrlIleThrAlaLeuProGlu 201
Db 547 GAGCAAAATCGCAATTCGCAACAGCAATGATGTCTATATATTAACGGCAAAAGCTGAG 606
Qy 202 TyTylybAlaMetAsnaerValAlaArgGlyIleValLybAlaAspValAlaGlnSerSer 221
Db 607 TATATbGATGATGAATGATGTGCACTGCAATTTGCAAAAGCGAGTGGCGAGAGCAGC 666
Qy 222 TyrglyLeuTyrglyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
Db 667 TACGGGTGTGATGCAAGAGACAGATCGTAGCGGTTCGACAGAGGCTTGATACAGT 726
Qy 242 ArgAsnaerSerSerMetHnIleGlnAlaPheArgGlyLybIleThrAlaLeuTyrlAlaLeu 261
Db 727 CGCAATGACAGTTCATGATGATGACCTTCGGGGGAAAATTAATCGCAATTATATGCAATG 786
Qy 262 GLyArgThrAsnaenAlaAsnaerThrAsnGlyHnIleGlyThrHnIleValAlaGlySerVal 281
Db 787 GGACGAGAGGAATAAAGCCAAATGATGATGCTCATGCTACCGCATGCTGGCTGCCCTCA 846
Qy 282 LeuGlyAsnGlyAlaThrAsnlybGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
Db 847 TTAGAAAACGGCTCACTTAATAAGAAATGGCGCTCAGGCGAAATCTAGTCTTCCAAATC 906
Qy 302 IleMetAspSerSerGlyLyLeuGlyLeuProSerAsnLeuGlnThrLeuPheSer 321
Db 907 ATCATGATAGCGGTGGGGGACTTGGAGGACTACCTTGGAATCTGCAAAACCTTATTCAGC 966
Qy 322 GlnAlaPheSerAlaGlyAlaArgIleHnIleThrAsnSerTrpGlyAlaAlaValAsnGly 341
Db 967 CAAGCATTCAGTGTGGTGGCCAGAAATTCATCAAACTCTGGGGGAGCGAGTGAATGGG 1026
Qy 342 AlaTyThrThrAspSerArgAsnaValaAspAspTyValArgLybAsnaSerMetThrIle 361
Db 1027 GCTTACACACAGATTCAGAAATGTGATGACTATGTGCGGAAAAGATATGCGATC 1086
Qy 362 LeuPheAlaAlaGlyAsnGlyLysProAsnGlyGlyThrIleSerAlaProGlyTyThrAla 381
Db 1087 CTTTTCCGTGCGGGAATGAGAGACGAAACGGCGGAACATACATGACGACCGGCAAGCT 1146
Qy 382 LybAsnaAlaIleThrValGlyAlaThrGlyLysnLeuAspProSerPheGlySerTyrla 401
Db 1147 AAAAATGCAATAACGTGGAGCTTCGGAACAACTCCCGCCAAAGCTTCGGGTCTTATCG 1206
Qy 402 AspaerTleAsnHnIleValAlaGlnPheSerSerArgLyProThrLybAsnGlyValArgIle 421
Db 1207 GACATATACACAGATGAGCAGATGCTTCTTCACTGAGCGACGACAAAGATGAGAGGATC 1266
Qy 422 LysProAspValMetAlaProGlyTyThrTyrlleuSerAlaArgSerSerLeuAlaPro 441
```

```
Db 1267 AAACCGATGTCATGGACCGGGAAAGCTTCATCATACAGCAAGATCTTCTTGACCG 1326
Qy 442 AspSerSerPheTrpAlaAsnHnIleAspSerTybTyrlAlaTyMetGlyGlyThrSerMet 461
Db 1327 GATTCCTCTTCCTGGGGAAACATAGACGTAATAATGCATACATGAGGTGGAAACGTCCATG 1386
Qy 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuHnrglybHnIlePheValLybAsnaArg 481
Db 1387 GCTTACACCGATCTGCTGGAACCTGGACACGCTTGTGAGCATTTTGTGAAAACAG 1446
Qy 482 GlyIleThrProLyProSerLeuLybAlaAlaLeuIleAlaGlyAlaAlaAspVal 501
Db 1447 GGCAATCACACAAAGCCCTTCTATTAAGCGGCACTGATTCGGGAGCAGACATC 1506
Qy 502 GlyLeuGlyTyTyProAsnGlyAsnGlnGlyTyTrpGlyArgValThrLeuAspLybSerLeu 521
Db 1507 GGCTTGAGCTTACCGGAAACGGTAAACGAGATGGGGAAGATGACATTTGAGTAAATCCCTA 1566
Qy 522 AsnValAlaTyValaAsnGlySerSerAlaLeuSerThrSerGlnLybAlaThrTyThr 541
Db 1567 AACGTTGCTATGTGAACGAGTCCAGTTCCTATTCACACGCAAAAGCACTACTCG 1626
Qy 542 PheThrAlaThrAlaGlyLybProLeuLybIleSerLeuValTrpSerAspAlaProAla 561
Db 1627 TTACTGCTACTGCGGCAAGCCTTTGAAGATCTCCCTGTATGTGCTGATGCCCTGCG 1686
Qy 562 SerThrThrAlaSerValThrLeuValAsnaPheLeuValIleThrAlaProAsn 581
Db 1687 AGCAACAAGTGTTCGTAAGCCTTGCAATGACCTTGACCTTGATACCTGCTCCAAAT 1746
Qy 582 GlyThrArgTyValGlyAsnaerPheSerAlaProPheAsnaenAntTrpAspGlyArg 601
Db 1747 GGCAACAATATGTGGAATATGACTTATCTTCGCAATACATATTAATCTGGATGGCCG 1806
Qy 602 AsnaenValGlnaenValPheIleAsnSerProGlnSerGlyThrTyThrIleGluVal 621
Db 1807 AATAAGCTAGAAATGATTAATTAATGCGCCAAAGCGGAGGTATACATTAAGAGTA 1866
Qy 622 GlnAlaTyThrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGCTTATTAACGTATCCGTTGACACACAGACCTTCTCGTTGGCAATTTGAAAC 1920

RESULT 12
LOCUS AR069954 3003 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 41 from patent US 5891701.
ACCESSION AR069954
VERSION AR069954.1 GI:7220842
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
1. 3003
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.75e-143 Length: 3003
Score: 2799.00 Matches: 541
Percent Similarity: 92.4% Conservative: 52
Best Local Similarity: 84.3% Mismatches: 45
Query Match: 85.7% Indels: 4
DB: 6 Gaps: 2

US-10-784-870-4 (1-639) x AR069954 (1-3003)
Qy 1 MetArglyb-----LybLybValPheLeuSerValLeuSerAlaAlaAlaIleLeu 17
```

```

Db      846 ATGAGAAAGAAAGATCGAAGAGGGTTTATTCGGTTTATTCAGTTGTCGACATATG 905
Qy      18 SerThValAlaLeuAenAnPProSerAlaGlyAspAlaArgThrPheAspLeuAspPhe 37
Db      906 TCTTCGTGCTTAAAGCAGCTCTTCTTCAATATGCGGCGAAACAATTTGAATTTGAGCTTT 965
Qy      38 LysGlyIleGlnThrThrThrAspValSerGlyPheSerIleGlnArgIleGlnThrGlyAla 57
Db      966 AAGGGATGAGACACTTACGCTAGAGAGCGCCACCAAGCAAGGAAAGGAAAGCGGAGAG 1025
Qy      58 AlaAlaPheLeuValGluSerGluAenValIleLeuLeuIleGlyLeuLeuIleGlyLeu 77
Db      1026 GCATCTTTCTTGTAAACCTGAAAATGCGAAAATCCCAAGAGATTCACAAAGAAACCTA 1085
Qy      78 GluThrValProAlaAenAnIleLeuHisIleValGlnPheAenGlyProIleLeuGlu 97
Db      1086 GAAGTAGTTCAGCGGATTAACAAGCTATATATGCTTCAATTTGACGGAACCTATTTTAG 1145
Qy      98 GluThrLysGlnLysLeuGluIleThrThrGlyAlaIleGlyIleLeuAspPheProAspPhe 117
Db      1146 GAAACCAACTTCACTACGAGAAACGAGGAGCAAAATTCGATTAACAACAGATTAAC 1205
Qy      118 AlaTyrlIleValGluIleGluGluGlyAspValGlnSerLysValArgSerIleGluHisVal 137
Db      1206 GCTTATATTTGTGAATATGATGCGGATGTAAAGCCGTAACCTAACCCATTTGCCATTTG 1265
Qy      138 GluSerValGluProGlyLeuProGlyPhePheIleAspProGlnLeuPheThrLysGly 157
Db      1266 GAATCGGTGAAACCAATTTACCTTATATTAATGACCCGCAATTTATTTCAAGAGA 1325
Qy      158 AlaSerThrLeuValIleValAlaLeuAlaLeuAspThrLysGlnAenAnIleGluValGln 177
Db      1326 GCTTCGAATATGTAAACAGTAGCTTTGATTAAGAAAGCAAGAAAGAAAGTAACGT 1385
Qy      178 LeuArgGlyIleGluGluIleAlaGlnTyrlValAlaSerAspAspValHisTyrlIleThr 197
Db      1386 TTAAGAGATTGGAACAATTTGCCCAATACCGCAACAATATGATGATTTATTCGTAAAC 1445
Qy      198 AlaLysProGluTyrlLysValLysAspAspValAlaArgGlyIleValIleValAlaAspVal 217
Db      1446 CCAAGGCCGAAATACCAAGTTTGAATGACGTGCGCCGTGCGCATTTGTGAAGACAGACGTC 1505
Qy      218 AlaGlnSerSerTyrlLysTyrlGlyGlnGlyGlnIleValAlaValAlaAspThrGly 237
Db      1506 GCACAAAATTAACCTTTGGCTTATATGGAACAAGACAATTTGTAGCAAGTTCTGATCTGCG 1565
Qy      238 LeuAspThrGlyArgAspAspSerSerSerMetHisGluAlaPheArgGlyLysIleThrAla 257
Db      1566 CTTGATACAGAAAGAAATGACAGTTGATGATGAAGCAATTCGCGGTAAAGATTACCGCA 1625
Qy      258 LeuTyrlAlaLeuGlyArgThrAspAspAlaAspAspThrAsnGlyHisGlyThrHisVal 277
Db      1626 CTATATGCACTGGGCGAAGCAATACGCAATGATCCAAATGGAATGGAATGCCATGTT 1685
Qy      278 AlaGlySerValLeuGlyAenGlyAlaThrAsnLysGlyMetAlaProGlnAlaAenLeu 297
Db      1686 GCTGATCTGTGTAGGAAT--GCTACAAATTAAGGATGCGACCGCAAGCCATTTCA 1742
Qy      298 ValPheGlnSerIleMetAspSerSerGlyLysGlyIleLeuProSerAenLeuGln 317
Db      1743 GCTCTTCAATCTATTTGATATGATGATGAGGGCTGGAGAGACTCTCTATTTACCA 1802
Qy      318 ThrLeuPheSerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTrpGlyAla 337
Db      1803 ACATTTCTCAGTCAAGCATATAGTCTGAGCGAAGATTCATACGAATTCATGGGGGCT 1862
Qy      338 AlaValAsnGlyAlaTyrlThrThrAspSerArgAspValAspAspTyrlValArgLysAen 357
Db      1863 CCAAGTAACGGGTGCTTATGCAAGACTCTCGAAAGTGAATATATATGAGAAATAAT 1922
Qy      358 AspMetThrIleLeuPheAlaAlaGlyAenGluArgProAsnGlyGlyThrIleSerAla 377

```

```

Db      1923 GATATGACGATTCCTTTTGGCGCGGAATGAGGGAACAAGTAGCGGTAACAATCACTGCA 1982
Qy      378 ProGlyThrAlaLysAenAlaIleThrValGlyAlaThrGluAenLeuArgProSerPhe 397
Db      1983 CCAGAAACAGCAAAATGCGATTAACGTTGGGCGAACCAAGAAACCTACGTCACAGCTTC 2042
Qy      398 GlySerThrAlaAspAsnIleAenHisValAlaGlnPheSerSerGlyProThrLys 417
Db      2043 GAGTCTTAAGCGGATATATTAACAGATTTGCTCAATTTCTTCAAGAGGCTCTACTAGA 2102
Qy      418 AspGlyArgIleLysProAspValMetAlaProGlyThrTyrlIleLeuSerAlaArgSer 437
Db      2103 GATGACGATTTAAGCGGACGTCATGAGCACAGAGTACGATATTTCTCTGCTAATCA 2162
Qy      438 SerLeuAlaProAspSerSerPheTrpAlaAenHisAspSerIleTyrlAlaTyrlMetGly 457
Db      2163 TCATTAGCTCCAGATTCCTCATTTGCGGCAACCAATGATGTAATATGCTCAAGGGT 2222
Qy      458 GlyThrSerMetAlaThrProIleValAlaGlyAspValAlaGlnLeuArgGlyHisPhe 477
Db      2223 GGTACTTCTATGCTATCTCAATTTAGAGAGTGAATGTTGCACATTTAAGGAGCATTTT 2282
Qy      478 ValLysAsnArgGlyIleThrProLysProSerLeuLeuValAlaLeuIleAlaGly 497
Db      2283 GTGAAAATAAGAGGCTAATCTCTAAGCTTCCTTTAAAGCTCCTTAATTCAGAGT 2342
Qy      498 AlaAlaAspValGlyLeuGlyTyrlProAenGlyAenGlnGlyTyrlValArgValThrLeu 517
Db      2343 GCTGCGGATGTTGACTTGGCTTTCCAAATGTGTAACCAAGATGGGGAAGATGACGTTA 2402
Qy      518 AspLysSerLeuAsnValAlaTyrlValAenGlyUserSerAlaLeuSerThrSerGlnLys 537
Db      2403 GATTAATCCCTTAATGTGCAATTTGATGAAGAAACAGGCCCTTTATCAACAATCAAAA 2462
Qy      538 AlaThrTyrlThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTrpSer 557
Db      2463 GCAACATATTCCTTTACGCTCAAGCTGTAACCTTAAATATCATCTGTTGGTCA 2522
Qy      558 AspAlaProAlaSerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIle 577
Db      2523 GATGCACAGTAGACGACGAGCACTACATCACTTAAGTAATGATTAACATTAAGTAAATC 2582
Qy      578 ThrAlaProAsnGlyThrArgTyrlValGlyAenAspPheSerAlaProPheAspAenAn 597
Db      2583 ACTGCACCAATGGAATCTTAACGTGCGAAATGACTTTACGACACCGTATATTAACAAT 2642
Qy      598 TrpAspGlyArgAspAsnValGluAenValPheIleAsnSerProGlnSerGlyThrTyrl 617
Db      2643 TGGGATGGCAAAACAAGTGAATATGTTATCTCATAGCTCTCAAGCGGACGATAT 2702
Qy      618 ThrIleGluValGlnAlaTyrlAsnValProValGlyProGlnAsnPheSerLeuAlaIle 637
Db      2703 ACAGTGAAGTCAAGCTTACATGATCAAGTAAAGTCCGCAAACTTTCTTTAGGATTT 2762
Qy      638 ValAen 639
Db      2763 GTACAT 2768

```

RESULT 13
 LOCUS BD062155 3003 bp DNA linear PAT 27-AUG-2002
 DEFINITION Nucleic acids encoding a polypeptide having protease activity.
 ACCESSION BD062155
 VERSION BD062155.1 GI:22607760
 KEYWORDS JP 2001514529-A/39.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 3003)
 REFERENCE Sloana A. and Christianson L.
 Nucleic acids encoding a polypeptide having protease activity

JOURNAL Patient: JP 2001514529-A 39 11-SEP-2001;
NOVO NORDISK BIOTECH INC
COMMENT PN JP 2001514529-A/39
PD 11-SEP-2001
PF 09-JUN-1998 US 1999503145
PR 12-JUN-1997 US 08/873479
PI ALAN SIOWA, LYNN CHRISTIANSON
PC C12N15/57, C12N15/75, C12N9/54, C12K14/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key: Location/Qualifiers.
FEATURES Location/Qualifiers.
source 1..3003
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1,75e-143 Length: 3003
Score: 2799.00 Matches: 541
Percent Similarity: 92.44 Conservative: 52
Best Local Similarity: 84.34 Mismatches: 45
Query Match: 85.74 Indels: 4
Gaps: 2
US-10-784-870-4 (1-639) x BD062155 (1-3003)
QY 1 MecArgLyS-----LyLyLyValPheLeuSerValLeuSerAlaAlaAlaLeu 17
DB 846 ATGAGAAAGAAAGATCGAAGAGGGTTTTTATCGTTTATCGATTGCTCAATATG 905
QY 18 SerThrValAlaLeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPhe 37
DB 906 TCTTCTGTGCTTAAAGGCTCTTCTATTTGGCGCGAACAATTTGAATTCGACCTT 965
QY 38 LysGlyIleGlnThrThrAspValSerGlyPheSerIleGlnGlnThrGlyAla 57
DB 966 AAGGGATAGAGACCTTACGCTAGAGAAAGCTGCCACCAAGCAAGAAAGCGGAAAG 1025
QY 58 AlaAlaPheLeuValGlySerGlyAsnValLysLeuLeuLysGlyLeuLeuLysLeu 77
DB 1026 GCATCTTTCTTGTAACTCTGAAATATGTAATAATCCCAAGAGATTCAAAGAAACTA 1085
QY 78 GlyThrValProAlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleuGln 97
DB 1086 GAAGTAGTTCCAGCGGATTAACAAGCTATATATCGTTCAATTTGAGCGACCTATTTAAG 1145
QY 98 GlyThrLysGlnLysLeuGlnThrThrGlyAlaLysIleLeuAspThrIleProAspThr 117
DB 1146 GAAGCGCACTTCAACAGAGAAAGCGGAGCGAAATTCCTGATTACATCCAGATTAC 1205
QY 118 AlaTyrIleValGlyTyrGlnGlyAspValGlnSerLysValArgSerIleGlnHisVal 137
DB 1206 GCTTATATGTCGATATGATGATGGGATGTAAGCGCGTAACGCAATTCGCATTTG 1265
QY 138 GlySerValGlyProGlyLeuProLysTyrLysIleAspProGlnLeuPheThrLysGly 157
DB 1266 GAATCGGTGAACCATATTACCTTATATTAATAAGACCCGCAATTTATTTCCAGAGGA 1325
QY 158 AlaSerThrLeuValAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGlyValGln 177
DB 1326 GCTTCTGAATTAAGTAACGTACTTAAATTAAGCAAGAAAGTAAGAGATGCT 1385
QY 178 LeuArgGlyIleGlnGlnIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThr 197
DB 1386 TTAAGAGATTGGAACAATTCGCCAATACCGCAACAATTAATGATGTTATACGTAAC 1445
QY 198 AlaLysProGlyTyrLysValLysAsnAspValAlaArgGlyTyrIleValLysAlaAspVal 217
DB 1446 CCAAGCGCTGAATACGAAGTTTGAATGACGTGGCCGTGGCATTTGTGAAGCAACGTC 1505
QY 218 AlaGlnSerSerTyrGlyLeuTyrGlyGlnGlnIleValAlaValAlaAspThrGly 237

DB 1506 GCAAAATAATTACTTTGGCTTATATGACAGACAGATTGTAAGCATTTGCTGATATCGG 1565
QY 238 LeuAspThrGlyArgAsnAspSerSerMetHisGlyLysAlaPheArgGlyLysIleThrAla 257
DB 1566 CTTGATACAGAGAAATATGACATTCATGATGATAGCACTCCCGGTAATTAACCGCA 1625
QY 258 LeuTyrAlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisVal 277
DB 1626 CTATATGACCTGGCGACAAACCAATACGCCAATATATCCAAATGACATGAAACCAAGTT 1685
QY 278 AlaGlySerValLeuGlyAsnGlyValaThrAsnLysGlyMetAlaProGlnAlaAsnLeu 297
DB 1686 GCTGATTCGTGTTAAGAAAT---GCTTAATAAAGGATGGCACCGCAAGCCAACTTA 1742
QY 298 ValPheGlnSerIleMetLysSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGln 317
DB 1743 GTCTTCAATCTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1802
QY 318 ThrLeuPheSerGlnAlaPheSerAlaGlyValaArgIleHisThrAsnSerTyrGlyAla 337
DB 1803 ACATTATTCAGTCAAGCATATATGCTGAGCGGAATTCATTAAGAAATTCATGGGGCT 1862
QY 338 AlaValAsnGlyAlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsn 357
DB 1863 CCAATTAACGGTGCCTATACAGACAGCTCTGAAATGTTATGATGATGATGATGATGAT 1922
QY 358 AspMetThrIleLeuPheAlaAlaGlyAsnGlyAsnGlyArgProAsnGlyTyrThrIleSerAla 377
DB 1923 GATATGACGATTTCTTTTGGCGCGAAATATGAGGACCAAGTACCGGTAACATCAGTCA 1982
QY 378 ProGlyThrAlaLysAsnAlaIleThrValGlyValaThrGlyAsnLeuAspProSerPhe 397
DB 1983 CCAAGAACAGAAAAAATGCAATTAACAGTTGGGCAACCGAAACCTACGCTCAAGCTTC 2042
QY 398 GlySerTyrAlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLys 417
DB 2043 GCATCTTATGCGGATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2102
QY 418 AspGlyArgIleLysProAspValMetAlaProGlyTyrTyrIleLeuSerAlaArgSer 437
DB 2103 GATGACGATATTAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2162
QY 438 SerLeuAlaProAspSerSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMetGly 457
DB 2163 TCATTAGCTCAAGATTCCTCAATTCGGGCAAAACATATATTAATTAATTAATTAATTAAT 2222
QY 458 GlyThrSerMetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlnHisPhe 477
DB 2223 GGTACTTCTATGCTACTCCCAATTTGACAGTATATTTGCAACAATTAAGGACATTTT 2282
QY 478 ValLysAsnArgGlyIleThrProLysProSerLeuLeuLysValAlaLeuIleAlaGly 497
DB 2283 GTGAAATAATAGAGGGTAACCTTAAGCTTCCTTTTAAAGCTGCTTAAATTCAGAGT 2342
QY 498 AlaAlaAspValGlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrProGlyArgValThrLeu 517
DB 2343 GCTCGGATGTGGACCTGCTTCCAAATGTGAACCAAGATGGGGAAGGTAACGTGA 2402
QY 518 AspLysSerLeuAsnValAlaTyrValaAsnGlySerSerAlaLeuSerThrSerGlnLys 537
DB 2403 GATTAATCCCTTAATATGTCGATTTGTGAATGAACAGACCTTTATCAACAAGTCAAA 2462
QY 538 AlaThrTyrThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSer 557
DB 2463 GCAACATATTCGTTTACGGCTCAAGCTGTGTAAACCTTAATAAATATCACTGTTTGGCA 2522
QY 558 AspAlaProAlaSerThrThrAlaSerValThrLeuValaAsnAspLeuAspLeuValIle 577
DB 2523 GATCAACAGGTATACAGACGCAATCACTTAATTAAGATGATTAAGATCTTGATATC 2582
QY 578 ThrAlaProAsnGlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsn 597

Db 2583 ACTGCACCAATGGAATAATACGTCCGGAATGACTTTACAGACCGGTATGATAACAA 2642

Qy 598 TTPAEPGLYARGAsenValGluAsnValPheIleAsnSerProGlnSerGlyThrTy 617

Db 2643 TGGATGGCGAAGAACGCGGAAATGCTTATCAATGCTCTCAAGCGGAAGTAT 2702

Qy 618 ThrIleGluValGlnAlaTyraAsnValProValGlyProGlnAsnSerIleuAlaIle 637

Db 2703 ACAGTGAAGTGCAGGCTTACATATGACAGTAAGTCCGCAACCTTTCTTTAGCGATT 2762

Qy 638 ValAsn 639

Db 2763 GTACAT 2768

RESULT 14

LOCUS E03808 2218 bp DNA linear PAT 29-SEP-1997

DEFINITION DNA encoding alkaline protease Ya.

ACCESSION E03808

VERSION E03808.1 GI:2172022

KEYWORDS JP 1992197182-A/1.

SOURCE Bacillus sp.

ORGANISM Bacillus sp.

REFERENCE 1 (bases 1 to 2218)

AUTHORS Tobe,S., Odera,M. and Asai,Y.

TITLE DNA CODING ALKALINE PROTEASE YA ENZYME AND PRODUCTION OF ALKALINE

JOURNAL PROTEASE YA USING THE DNA

Patent: JP 1992197182-A 1 16-JUL-1992;

LION CORP

OS Bacillus sp.

PN JP 1992197182-A/1

PD 16-JUL-1992

PF 28-NOV-1990 JP 19903271110

PI TOBE SEIICHI, ODERA MOTOMASU, ASAI YOSHIO

PC C12N5/57,C11D3/386,C12N9/54,(C12N5/57,C12R1:07),(C12N9/54,

PC C12R1:07);

CC strandedness: Double;

CC topology: linear;

CC hypothetical: No;

CC anti-sense: No;

CC *source: strain:Y strain;

CC *source: clone=PB8A;

CC Key Location/Qualifiers

FT CDS 218..2125

FT 1. /product=precursor of alkaline protease Ya'

FT sig_peptide 218..823

FT mat_peptide 824..2122

FT 1. /product=precursor of alkaline protease Ya'.

FEATURES

source Location/Qualifiers

1..2218

/organism="Bacillus sp."

/mol_type="genomic DNA"

/db_xref="taxon:1409"

ORIGIN

Alignment Scores:

Pred. No.: 2,276-127 Length: 2218

Score: 2500.50 Matches: 480

Percent Similarity: 87.94 Conservative: 80

Best Local Similarity: 75.44 Mismatches: 72

Query Match: 76.54 Indels: 5

DB: 6 Gaps: 4

US-10-784-870-4 (1-639) x E03808 (1-2218)

Qy 3 LysIleValSerValPheIleSerValIleuSerAlaAlaIleIleuSerThrValAlaIleu 22

Db 227 AAAAAAAGATGATGCTATGATGCTTCGCTGCGATCTTAAGCGCTGATTAATGGTT 286

Qy 23 AsnAsnProSerAlaGlyAspAlaArgThrPheAspIleuAspPheIleGlyIleGlnThr 42

Db 23

Db 287 AGTTCCACCACTAGTGG-----GCAATTTTCAAGTAATTTTAATGCTGGAAGT 340

Qy 43 ThrThrAspValSerGlyPheSerIleGlnArgIleThrGlyAlaAlaIlePheIleuVal 62

Db 341 TTAGAANAATGCTAGT---TTGGTTAAACCATTAAGTACGCGGTGAGCATCTTTCTAGTA 397

Qy 63 GlnSerGluAsnValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 82

Db 398 GATACGGAANAATTAATTAATTTCTTAAGGTATTCAAAGAACTGAGACATGACAGAAAG 457

Qy 83 AsnAsnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 102

Db 458 GATAAGCAACTCTACATGCTACATTTACTGCAACCAATTTCTGAGAGAAAGGAAAGCA 517

Qy 103 LeuGluThrThrGlyAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 122

Db 518 TTAGAGTCTCTAGAGATGATGATTCAGATTATGTTCCAGATTAGCTTTATTTGTTGAG 577

Qy 123 TyrGluGluAspValGlnSerIleuValArgSerIleuIleuIleuIleuIleuIleuIleu 142

Db 578 TATAGTGTGCTACCAAAAAAT---ATAAGTACTTTACATTCTGTTGAGAACGTACACCA 634

Qy 143 TyrIleuProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 162

Db 635 TTTTTCATCATTAATTAATTAATGATCTGACCTTTTAAGCAAGAGTCTTCCAGCTTGT 694

Qy 163 LysAlaIleuAlaIleuAspThrIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 182

Db 695 CAAAGCGTTATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 754

Qy 183 GluIleuAlaGlnIleuValAlaSerAsnAspValIleuIleuIleuIleuIleuIleuIleu 202

Db 755 GAGATCGTTCAATTAATGCTGCAAAATTAATGATGCTTTATTAATTAATTAATTAATTAAT 814

Qy 203 LysValIleuAsnAspValAlaArgGlyIleuValIleuValIleuValIleuValIleuVal 222

Db 815 GAGCTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 874

Qy 223 GlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 242

Db 875 GGATTAATTAATGCAAGGCTCAACTAGTTCAGTAAGGCAACAGGCTTGAATACAGTGGT 934

Qy 243 AsnAspSerSerMetHisGluAlaPheArgGlyIleuIleuIleuIleuIleuIleuIleuIleu 262

Db 935 AACGATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 994

Qy 263 ArgThrAsnAsnAlaAsnAspThrAsnGlyIleuIleuIleuIleuIleuIleuIleuIleu 282

Db 995 AGAATTAATTAATGCAAGTATCCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 1054

Qy 283 GlyAsnGlyAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 302

Db 1055 GGTAAAT---GCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1111

Qy 303 MetAspSerSerGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 322

Db 1112 ATGATTAATGACGAGGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1171

Qy 323 AlaPheSerAlaGlyAlaArgIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 342

Db 1172 GCTTGAAATGCTGAGCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1231

Qy 343 TyrThrThrAspSerArgAsnValAspAspTyrValArgIleuIleuIleuIleuIleuIleu 362

Db 1232 TACACTGCTAACTCGAAGCAAGTGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1291

Qy 363 PheAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 382

Db 1292 TTTCACAGCTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1351

Qy 383 AsnAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 402

Db 1352 AATGCTATTACGATCGGCGCAACGAAATTAATGATTAATTAATTAATTAATTAATTAATTAAT 1411

Db 721 TGGGCGAACCATGACAGTAATAATGCAATACATGGGTGGAACTGCATGGCTACACCGATC 780
Qy 466 ValAlaGlyAsnValAlaGlnLeuArgGlnHisPheValIleAsnArgGlyIleThrPro 485
Db 781 GTTGTGGAAACGTGGCAACGTTCTGAGCATTTGTGAAAAACAGAGCATCACACCA 840
Qy 486 LysProSerLeuLeuLysAlaLeuIleAlaGlyValAlaAspValGlyLeuGlyTyr 505
Db 841 AGCCCTTCTCTATTAAAGGGGCACTGATGGCGGTGCACCTGACATCGGCTTGGCTAC 900
Qy 506 ProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeuAsnValAlaTyr 525
Db 901 CCGAACGGTAACCAAGATGGGGACGAGTGCATTGGATTAATCCCTGAACGTTGCTTAT 960
Qy 526 ValAsnGlySerSerLeuSerThrSerGlnLysAlaThrTyrThrPheThrAlaThr 545
Db 961 GTGAACGAGTCCAGTCTCTATCCACACGCAAAAAGCGACGTACTCGTTACTGCTACT 1020
Qy 546 AlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAlaSerThrAla 565
Db 1021 GCCGGCAACCTTTGAAATCTCCCTGTATGCTGATGCCCTGCGAGCACACTGCT 1080
Qy 566 SerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyThrArgTyr 585
Db 1081 TCCGTAAACGCTTGTCAATGATCTGACCTTGTCAATACCGCTCAATGGCACACAGTAT 1140
Qy 586 ValGlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGlyArgAsnAsnValGlu 605
Db 1141 GTAGGAATGACTTATCTTCGCCATCAATGATTAACCTGGGATGGCCGCAATAACTAGAA 1200
Qy 606 AsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluValGlnAlaTyrAsn 625
Db 1201 AATGATTTTATTAATCAACCAACAGCGGACGTTATACAAATGAGGTACAGGCTTATAAC 1260
Qy 626 ValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1261 GTACCGGTTGACCAACACACTTCTCTGTGGCAATTGTGAAT 1302

Search completed: April 8, 2006, 05:26:55
Job time : 7515 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 8, 2006, 03:09:11 ; Search time 1640 Seconds
(without alignments)
2596.792 Million cell updates/sec

Title: US-10-784-870-4
Perfect score: 3267
Sequence: 1 MKKKVFLSLVLSAAAILSTV.....EVOAVNVVPQNPPLATVN 639

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame.p2n.model -DEV=xlp
-O=/abs/ABSSWEB-POOL/US10784870/runat_07042006_143414_27447/app_query.fasta_1
-DB=N Geneseg -OPMT=faetap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=pct -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs802p
-USER=US10784870@cgc1_1_1929 -runat_07042006_143414_27447 -NCPU=6 -ICPU=3
-NO WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseg_21.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2004as.*
12: geneseqn2004bs.*
13: geneseqn2004cs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3267	100.0	1920	2	AAK37277
2	3072	94.0	1923	2	AAK37279
3	3066	93.8	1923	2	AAK37278
4	3066	93.8	1923	12	ADM40772

Result No.	Score	Query Match	Length	DB ID	Description
5	3066	93.8	1923	13	ADSL4428
6	3066	93.8	1923	13	ADT49605
7	2799	85.7	3003	2	AAV82382
8	2183	66.8	1305	11	AAV33779
9	2183	1305	12	ADL25803	
10	1990.5	60.9	1299	2	AAQ27516
11	477	14.6	1977	2	AAH85667
12	477	14.6	1977	2	AAH85667
13	443	13.6	1962	2	AAH85665
14	443	13.6	1962	2	AAH85665
15	443	13.6	1962	2	AAH85665
16	443	13.6	1962	2	AAH85665
17	435	13.3	110000	12	ADN46845_14
18	435	13.3	110000	12	ADN47591_05
19	435	13.3	110000	12	ADN47591_06
20	435	13.3	110000	12	ADN46123_14
21	435	13.3	110000	12	ADN47209_05
22	435	13.3	110000	12	ADN47209_06
23	435	13.3	110000	12	ADN46464_14
24	435	13.3	110000	12	ADN47960_05
25	432	13.2	110000	12	ADN47960_06
26	417.5	12.8	1336	2	AAH85669
27	417.5	12.8	1336	2	AAH85669
28	392	12.0	2121	4	ABL54900
29	382	11.7	3684	12	ADM99122
30	357.5	10.9	2539	2	AAH61454
31	357.5	10.9	2539	2	AAH61455
32	345.5	10.6	1656	12	ADM99240
33	344.5	10.5	135638	10	ABX34289
34	344.5	10.5	104096	13	ADX56092
35	342.5	10.5	1329	6	ABK74643
36	339.5	10.4	1332	12	ADM99124
37	326	10.0	2835	2	AAH8141
38	326	10.0	4765	2	AAH8132
39	326	10.0	4765	2	AAH85670
40	326	10.0	4765	2	AAH85670
41	325.5	10.0	89378	12	ADN47591_20
42	325.5	10.0	89378	12	ADN47209_20
43	325.5	10.0	89378	12	ADN47960_20
44	325.5	10.0	110000	12	ADN46845_00
45	325.5	10.0	110000	12	ADN46123_00

ALIGNMENTS

RESULT 1
AAK37277
ID AAK37277 standard; DNA; 1920 BP.
XX
AC AAK37277;
XX
DT 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
DE Bacillus alkaline protease encoding DNA.
XX
KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW washing composition; oxidising agent; ss.
XX
OS Bacillus sp.
XX
PN W09918218-A1.
XX
PD 15-APR-1999.
XX
PF 07-OCT-1998; 98WO-JP004528.
XX
PR 07-OCT-1997; 97JP-00274570.
XX
PA (KAO) KAO CORP.
XX
PI Takaiwa M, Okuda M, Sasaki K, Kubota H, Hitomi J, Kageyama Y,
PI Shikata S, Nomura M,

XX WPI; 1999-287736/27.
DR P-PSDB; AAY17087, AAY17089.
XX Alkali protease from *Bacillus* used in washing powders.
XX
PS Disclosure: Page 53-58; 71pp; Japanese.
XX The invention relates to alkaline proteases produced by strains of
CC *Bacillus*. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 1920 BP; 576 A; 411 C; 473 G; 460 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,736-243 Length: 1920
Score: 3267.00 Matches: 639
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-784-870-4 (1-639) x AAX37277 (1-1920)
QY 1 MetAglYbLYbLYbValPheLeuSerValLeuSerAlaAlaAlaLeuSerThrVal 20
DB 1 ATGAGAAAGAAAGAGAGTGTGTTTATCTGTTTATCACTGCGACCATCTCTGCGACTGT 60
QY 21 AlaLeuAaAaAaProSerAlaGlyAaPaLaargThrPheAspLeuAspPheGlySer 40
DB 61 GCATTAAACAAATCCCTCGGCTGTCATGCAAGACATTGTTGATCTGATTTTAAAGAAAT 120
QY 41 GlnThrThrThrAspValSerGlyPheSerGlyGlnArgGlnThrGlyAlaAlaAlaPhe 60
DB 61 GCATTAAACAAATCCCTCGGCTGTCATGCAAGACATTGTTGATCTGATTTTAAAGAAAT 120
QY 41 GlnThrThrThrAspValSerGlyPheSerGlyGlnArgGlnThrGlyAlaAlaAlaPhe 60
DB 121 CAAGAAACAAACGAGATGTCAGTGTTCCTCAAAACGACCAAAACGGCGCTCATTT 180
QY 61 LeuValAGuSerGluAaAaValIysLeuLeuGlyLeuLeuValLeuGluThrVal 80
DB 181 CTGGTGGAGTCTGAAGAAATGTGAACCTTCTTAAGAGATTCCTAAAGAACTTGAAACAGTA 240
QY 81 ProAlaAaAaAaAaValLeuHisIleValGlnPheAaGlyProIleLeuGluThrVal 100
DB 241 CCGGCAAAATTAATTAATCTCATATGTCTCAATTCATATGGCCCATTTTGAAGAAACAAA 300
QY 101 GlnIysLeuGluThrThrGlyAlaIysIleLeuAspTyrIleProAspTyrAlaTyrIle 120
DB 301 CAGAGCTAGAGAACAACTGAGCAAAAGATTCCTGACTACATCCCTGATTTATGCATATAT 360
QY 121 ValGluTyrGluGlyAaAaValGlnSerIysValIasSerIleGluHisValGluSerVal 140
DB 361 GTCCAGATGAGAGGAGATGTTCAATCAAAAGTCCCTCATTTGAAACAGTGAATCAAGT 420
QY 141 GluProTyrLeuProIysTyrIysIleAspProGlnLeuPheThrIysGlyValAspThr 160
DB 421 GAGCCATCTCTGCGAAATTAACAAATTAAGATCCCGAGCTTTTCAAAAAGCGCATGACG 480
QY 161 LeuValIysAlaLeuAlaLeuAspThrIysGlnAaAaGluValGlnLeuArgGly 180
DB 481 CTGGGAAAGCGTTGCGCTTGATCGAAGCAAAACATTAAGAAAGTGAATTAAGAGCG 540
QY 181 IleGluGluIleAlaGlnTyrValAlaSerAaAaValHisTyrIleThrAlaIysPro 200

DB 541 ATCGAGAAATCGCTCAGTAGTACAGCAAGATGACATTAATTAACGCAAGCCT 600
QY 201 GluTyrIysValMetAaAaAaAaValAlaArgGlyIleValIleValAlaAspValAlaGlnSer 220
DB 601 GAATTAAGAGTGAATGAATGATGTGCCAGAGTATTTGCAAAAGCGAGTGTGCACAGAGC 660
QY 221 SerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAaAspThrGlyLeuAspThr 240
DB 661 AGCTACGGTGTGTATGACAAAGCCAGATTTCTGACATGGCCAGTACAGTGAATGATCA 720
QY 241 GlyArgAa 260
DB 721 GGAAGAAACGACAGTTCATGATGAGACCTTCGCGGTAAATTAACAGCATATATCA 780
QY 261 LeuGlyArgThrAa 280
DB 781 CTGGGTCCGACGAATTAATGCAATGATCGAAGCGTCACTGATCCCATGCGACGCTTCCG 840
QY 281 ValLeuGlyAaGlyValAlaThrAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 300
DB 841 GTATTAGGAAATGGCGCAACAAATTAAGAAATGGCACTCAAGCGAATCTGATTTTCA 900
QY 301 SerIleAa 320
DB 901 TCCATCATGATAGACAGTGTGGCTTGAAGCTTGCCTTCCATCTGCAAACTTATTC 960
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIleHisThrAaAaAaAaAaAaAaAaAaAaAaAa 340
DB 961 AGCCAGATTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 1020
QY 341 GlyAlaTyrThrThrAspSerArgAaAaValAspAspTyrValArgIysAaAaAaPheThr 360
DB 1021 GGGGCTACAGACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 361 IleLeuPheAlaAlaGlyAa 380
DB 1081 ATTCTTTTGGGGCTGGGAATGAAGCGCAAGCGGCTACATCAAGTGCACCTGTATCG 1140
QY 381 AlaIysAaAaAaAaIleThrValGlyAlaThrGluAaAaAaAaAaAaAaAaAaAaAaAaAa 400
DB 1141 GCTAAAAACGCAATTAACAGTGTGCGCAACCAAAACCTGCTCAAGCTTCTGCTCAT 1200
QY 401 AlaAa 420
DB 1201 GAGATTAATTAATCAAGTGTGCAAGTTCCTTCCCTGCGCCGCAAAAGATGGCGCA 1260
QY 421 IleIysProAa 440
DB 1261 ATCAAGCTGATGTCATGGCGCCAGGACATATCAATTTATCAGCAAGATCTTCTTCA 1320
QY 441 ProAa 460
DB 1321 CCCGATTCCTCTTCTGCGCAATCATGACGAAATATGCTATATGGGTGGAACGCTCC 1380
QY 461 MetAlaThrProIleValAlaAlaGlyAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 480
DB 1381 ATGGCAACAGATTTGTGGGGAATTTCAAGCTCCGTAAGCATTTTGTGAATAAT 1440
QY 481 ArgGlyIleThrProIysProSerLeuLeuIysAlaAlaLeuIleAlaGlyAlaAlaAaAa 500
DB 1441 AGAGGAATCATCTCAAGCTTCCCTATTTGAAGAGCTTTGATTCAGAGGCTGTAT 1500
QY 501 ValGlyLeuGlyTyrProAa 520
DB 1501 GTTGATTTGGGTATCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1560
QY 521 LeuAa 540
DB 1561 TTGAACGTGCTATGTAAGAAATCAAGTCCCTATCACTAGCCAAAGCAAGCAATAT 1620
QY 541 ThrPheThrAlaThrAlaGlyIysProLeuIysIleSerLeuValIleProSerAspAlaPro 560
DB 1621 ACCTTTACTGACAGCGCGGCAAGCATTTGAATCTCCCTGATGTGCTCGAGTCCCT 1680

QY 561 AIASeThrThraIaSerValThrlleuValaAsnAspLeuAsnleuValleThraIaPro 580
DB 1681 GGAAGCACTACGCTCTCTGTAACCTGGTCAATGATTTGGATTCATTACAGCACCA 1740
QY 581 AaNgIyThrArgTyValGlyAsnAspPheSerAlaProPheAspAsnAsnTPAPGly 600
DB 1741 AAGGAAACAAGATATGTCGGGAATGACTTCTACGACACATTTGACATTAATCGGATGGC 1800
QY 601 ArgAsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrYThrIleGlu 620
DB 1801 CCCAATAACGTGAAAATATATTATTAAATCGCCCAAGTGAACATATACCATTTGAG 1860
QY 621 ValGlnAlaTyAsnValProValGlyProGlnAsnAspSerleuAlaIleValaAsn 639
DB 1861 GTGCAAGCATATATATGTCGGTGGACCAACAACTCTCTGTCGCAATGTGAAC 1917
RESULT 2
AAK37279 standard; DNA; 1923 BP.
AC AAK37279;
DT 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
DE Bacillus alkaline protease encoding DNA.
KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW washing composition; oxidising agent; ss.
OS Bacillus sp.
XX W09918218-A1.
XX 15-APR-1999.
XX 07-OCT-1998; Y98WO-UP004528.
XX PR 07-OCT-1997; 97JP-00274570.
XX PA (KAO) KAO CORP.
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
XX Shikata S, Nomura M;
XX WPI; 1999-287736/27.
XX DR P-PsDB; AAV17091.
XX PT Alkali protease from Bacillus used in washing powders.
XX PS Disclosure; Page 63-68; 71pp; Japanese.
XX CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidising agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
SQ Sequence 1923 BP; 578 A; 417 C; 474 G; 454 T; 0 U; 0 Other;
Alignment Scores: 4.62e-228 Length: 1923
Pred. No.: 3072.00 Matches: 596
Score:

Percent Similarity: 97.0% Conservative: 23
Best Local Similarity: 93.4% Mismatches: 19
Query Match: 94.0% Indels: 0
DB: 2 Gaps: 0
US-10-784-870-4 (1-639) x AAK37279 (1-1923)
QY 2 ArgGlyValValPheIleuSerValIleuSerAlaIleAlaIleIleuSerThrValAla 21
DB 7 AAGAAAGAAAGAGTGTATCTGTTTATCTGTTTATCAGTCGACGCAATTTGTGCACTGTGGG 66
QY 22 LeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheIleGlu 41
DB 67 TTAAGTAACTCAATCTGCAAGGTGTGCAAGAAATTTGAATCTGAATTTCAAGAAATTCAG 126
QY 42 ThrThrThrAspValSerGlyPheSerIleGlnArgIleThrGlyValAlaIlePheIle 61
DB 127 ACAACAACAGATGCTAAAGCTTCTTCAAGACAGGGGCAAGCTGTCTCTCTTTCTG 186
QY 62 ValGluSerGluAsnValIleLeuLeuIleValGlyLeuLeuValGlyLeuThrValPro 81
DB 187 GTGGAATCTGAAAATGTGAACCTCCCAAAAGGTTTGCAGAAAGAGCTTGAAACAGTCCCG 246
QY 82 AIAAsnAsnValLeuHisIleValGlnPheAsnGlyProIleLeuGlnGluThrIleGln 101
DB 247 GCAATTAATAAATCTCATATTAATCAATTAATCAATTAATTAATTAATTAATTAATTAATTA 306
QY 102 IyLeuGluThrThrGlyValAlaValIleLeuAspTyIleProAspTyIleValIleVal 121
DB 307 CAGCTGGAAAAACAGGGGCAAAAGTTCTCGACTACATACCTGATTAATGCTTACATTTGTC 366
QY 122 GluTyGluGlyAspValGlnSerIleValArgSerIleGluHisValGluSerValGlu 141
DB 367 GAGTATGAGGGCGAGTAAAGTCAAGCAACACATTCAGACAGTGGAAATCCGTGGAG 426
QY 142 ProTyIleuProLyTyIleIleAspProGlnIleuPheThrIleGlyValIleSerThrIleu 161
DB 427 CTTATTTGCGATTAATCAAGATTAATCAATTAATCAATTAATTAATTAATTAATTAATTAAT 486
QY 162 ValIyValAlaValAlaValAspThrIleGlnIleAsnAsnValGluValGluIleuArgGlyIle 181
DB 487 GTTAAAGAGAGTGGCGCTTGAATCAAGACAAAGAAATTAAGAGTGCATTAAGAGGCAATC 546
QY 182 GluGluIleAlaGlnTyValAlaIleSerAsnAspValHisTyIleThrAlaIleAspProGlu 201
DB 547 GAACAAATCGCAATTCGCAATTAAGCAATGATGTCTATATTAATTAATTAATTAATTAATTA 606
QY 202 TyIleValIleMetAsnAspValAlaArgGlyIleValIleValAlaAspValAlaGlnSerSer 221
DB 607 TATAAGGTGAATGAATGATGTGGCGGTGAATTTCAAGCGAGATGTGCTCAGAGCAGC 666
QY 222 TyrGlyLeuTyGlyGlnGlyGlnIleValAlaValAlaAspThrGlyIleuAspThrGly 241
DB 667 TACGGTGTCTATGACAAAGACAGATCTTAGCGGTTCGATACAGGGCTTGAATCAAGGT 726
QY 242 ArgAsnAspSerSerMetHisGluAlaPheArgGlyIleIleThrAlaIleuTyValAlaIleu 261
DB 727 CGCAATGACAGTTCATGATCAAGAGCTTCGCGGGGAAATTAATGACATTAATATCAATTC 786
QY 262 GlyArgThrAsnAsnAlaAsnAspThrArgGlyHisGlyThrHisValAlaGlySerVal 281
DB 787 GACCGAGCAAGATTAAGTCAATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 846
QY 282 LeuGlyAsnGlyAlaThrAsnValGlyMetAlaProGlnAlaAsnValAlaPheGlnSer 301
DB 847 TTAGAAACCGGCTCACTAAATTAAGATGCGCTTACAGGGAATCTAGTCTTCAATCT 906
QY 302 IleMetAspSerSerGlyIleuGlyIleuProSerAsnLeuGlnThrIleuPheSer 321
DB 907 ATCATGATTAACGCGGTGGGAGCTTGAAGACATCTTGAATCTGCAACCTTATTCAGC 966
QY 322 GlnIlePheSerAlaGlyValArgIleHisThrAsnSerTPGValAlaValAsnGly 341

967 CAGCATACAGTGTGTCGTCAGAAATTCATACAACTCCTGGGAGCAGCATGAAATGGG 1026
QY AATATTTTThrAspSerArgAsnValAspAspTyrValArgValAsnAspMetThrIle 361
DB 1027 GTTACACACAGATTTCCAGAAATGTGATGACTATGTGCGCAAAAATGATATGACGATC 1086
QY 362 LeuPhealAAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAla 381
DB 1087 CTTTCGCGCGCGGAAATGAAAGACCGAAGCGGAAACCATCATGTGCACAGCAGCT 1146
QY 382 LyAsnAlaIleThrValGlyValaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
DB 1147 AAAAATGCAATTAACAGTCCGAGCTACGGAACCTCCGCCCAAGCTTTGGGCTTATGCG 1206
QY 402 AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrIleAspGlyArgIle 421
DB 1207 GACAAATATCAACCAATGTGACAGATTCCTTCACGTGACCCACAAAGAGATGAGGATTC 1266
QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
DB 1267 AAACCGAATGTCAATGGCACCGGGAACGTTCAATCACTATCGCAAGATCTTCTTGCACCG 1326
QY 442 AspSerSerPheTyrAlaAsnHisAspSerIleTyrAlaTyrMetGlyGlyThrSerMet 461
DB 1327 GATTCCTCTCTCTGCGCGCAACCATGACAGTAATATGATACATAGGTTGGAACGTCATG 1386
QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPheValIleAsnArg 481
DB 1387 GCTACACGATGTGTGCGAAACGTGGACAGCTTCGAGACATTTGTGAAAACGA 1446
QY 482 GlyIleThrProLysProSerLeuLeuValAlaLeuIleAlaGlyAlaAlaAspVal 501
DB 1447 GGCATCACACCAAGACCTTCTCTATTAAGGCGGACATGATGGCGGTCAGCTACATC 1506
QY 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeu 521
DB 1507 GGCCTTGGGTACCCCAACGCTAACCAAGATGGGACGATGACATTTGATTAATCCCTCG 1566
QY 522 AsnValAlaTyrValAsnGluSerSerAlaLeuSerThrSerGlnValaThrTyrThr 541
DB 1567 AACGTGCTTATGTAAACGATCCAGTTCTCTATCCACCAAGCAAAAGCGATCACTCG 1626
QY 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAla 561
DB 1627 TTTATGCTACTGCGCGCAAGCTTTGAATAATCTCCCTGATGTGATGCTGCTCG 1686
QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
DB 1687 AGCAACAACCTGCTCCGTAAACGCTTGTCAATGATCTGGACCTTGTCAATACCCCTCAAT 1746
QY 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnThrAspGlyArg 601
DB 1747 GGCACACGATATGTAAGAAATACCTTACTTCGCCATTCATATGATACCTGGGATGGCGC 1806
QY 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
DB 1807 AATAACGTAAGAAATGATTTATTAATGACACCAAGCGGACCTTATACATTAAGAACTA 1866
QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
DB 1867 CAGGCTTATTAACGTACCGGTGGACCAAGAACTTCTGTTGGCAATGTGAAAT 1920
RESULT 3
AAK37278
ID AAK37278 standard; DNA; 1923 BP.
XX AAK37278;
XX AC
XX 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
DE Bacillus alkaline protease encoding DNA.

KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KM washing composition; oxidizing agent; 88.
OS Bacillus sp.
XX
XX WO910218-A1.
XX
PD 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX (KAO) KAO CORP.
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX
XX MPI; 1999-287736/27.
DR P-PsDB; AAY17088, AAY17090.
XX
XX Alkali protease from Bacillus used in washing powders.
PT
PS Disclosure; Page 58-63; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkali protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,35e-227 Length: 1923
Score: 3065.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: Gaps: 0
US-10-784-870-4 (1-639) x AAK37278 (1-1923)
QY 2 ArgLysIleValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrValAla 21
DB 7 AAGAAAGAAAAAGGTTTATCTGTTTATCACTCAGCGAATTTGTGCAGCTGGCG 66
QY 22 LeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
DB 67 TTAAGTATTCATCTGACGGGTGCAAGGAATTTATGATCTGCAATTTCAAGGAATTCAG 126
QY 42 ThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaAlaPheLeu 61
DB 127 ACAACAATGATGTAAGGTTTCTCCAGAGAGGGGAGACTGCTGCTGCTTTCTG 186
QY 62 ValGluSerGluAsnValIleValLeuLeuLysGlyLeuLeuLysIleValPro 81
DB 187 GTGGAATCTGAAATGTGAAACTCCCAAAAGTTTGCAGAAAGAGCTTGAACAGTCCCG 246
QY 82 AlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlnGluThrLysGln 101
DB 247 GCAATATATTAATCTCATATTATCAATTCATATGACCAATTTTGAAGAAACAAACAG 306
QY 102 LysLeuGlnThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121

```

Db      307  CAGCTGAAAAAACAAGGGGAAAGATTCTGACTACACTGATTATGCTTACATTGTC 366
      122  GUYTGTGUGLYAaPValGInSerLyVaLaArgSerLLeGIuHIVaLGIuSerVaLGIu 141
      367  GAGTAAAGAGGGCGATGTTAAAGTCAAGACAAGCAAGCAATTGAGCAGCTGGATCCGGGAG 426
      142  ProLYrLeuProLYrLeuPProGInLeuPheThrLyLeuLYrLeuLYrLeu 161
      427  CCTATTTCCTGATTAACAGAAATGATGCCAGCTTTTCACAAAAAGGGGATCAGAGCTT 486
      162  ValLYrVaLaLeuLYrLeuPProGInLeuPheThrLyLeuLYrLeuLYrLeu 181
      487  GTPAAAGAGTGGCGCTTGATACAAAGCAAGAAAATTAAGAGGTGCATTAAGAGGCATC 546
      182  GIUGLIILeLaGInLYrVaLaLaSerAaPValHIVaLYrLeuThraLYrLeuPProGIn 201
      547  GAACAATCGCAATTCGCAATTAAGCAATGATGTCATATATATATGCGCAAGGCTGAG 606
      202  TYrLYrVaLaMeaAaPValAaArgGYrLeuLYrVaLaLaSerVaLaAGInSerSer 221
      607  TATTAAGTATGATATGATGTCGCGGTGAATGTCAAAAGCGGATGTGCTCAGAGCAC 666
      222  TYrGYrLeuTYrGYrGInGInLYrVaLaLaVaLaLaAaPProGInLYrLeuPProGIn 241
      667  TACGGGTTGATGAGCAAGGACAGATCGTACGGCTTGCCGATACAGGCTTGATACAGGT 726
      242  ArgAaPProSerSerLeuHIVaLaAaPProGInLYrLeuLYrLeuLYrLeu 261
      727  CCCAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
      262  GUYrGTThrAaPValAaPProGInLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 281
      787  GGAAGGACCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
      282  LeuGInLYrVaLaLaAaPProGInLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 301
      847  TTAGGAAAGCGCTCCATTAATTAAGGAAATGGCGCTCAGGCGGATTAATCTTCCATCT 906
      302  ILeuAaPProSerSerLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 321
      907  ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
      322  GInLaAaPProSerSerLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 341
      967  CAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
      342  AlaLYrThrThraAaPProGInLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 361
      1027  GCTTACCAACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
      362  LeuPheLaLaLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 381
      1087  CTTTTCGCTGCGGGAATGAAGGACCGGAAACCATCATGATGATGATGATGATGATGATGAT 1146
      382  LYrAaPProSerSerLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 401
      1147  AAAAAGGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206
      402  AaPProSerSerLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 421
      1207  GACATATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
      422  LYrPProAaPValMeaLaPProGInLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 441
      1267  AAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
      442  AaPProSerSerLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 461
      1327  GATTCCTCTTTCGGGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386
      462  AlaThrProGInLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 481

```

```

Db      1387  GCTACACGATCGTGTGCTGAAAAAGTGGACAGACTTCTGTGAGCAATTTTGAAAAACAGA 1446
      482  GUYrLeuThraPProSerLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 501
      1447  GGCATCAACCAAAAGCTTCTATTAATAAGGGGACATGATGCGGTCACCTGACATC 1506
      502  GUYrLeuGInLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 521
      1507  GGCCTTGGCTTACCGGAAACGATTAACCAAGATGGGAGGAGTGCATGATTAATATCTCTG 1566
      522  AaPValAaLYrVaLaAaPProGInLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 541
      1567  AAAGTTCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626
      542  PheThraAaPProSerLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 561
      1627  TTTACTGCTTACCGCGGCAAGCTTGAATACTCCCTGATGATGATGATGATGATGATGATGAT 1686
      562  SerThraAaPProSerLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 581
      1687  AGCAACATGCTTCCGTAACGCTTGTCAATGATCTGACCTTGTCAATGATCTGACCTTCAAT 1746
      582  GUYrThraGYrVaLaGInLYrAaPProSerLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 601
      1747  GGCACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1806
      602  AaPValGInLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 621
      1807  AATTAAGTAAAGAAATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1866
      622  GInLaLYrAaPProGInLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeuLYrLeu 639
      1867  CAGGCTTATTAAGTACCGGTTGACCAAGACCTTCTCGTGGCAATGATGAT 1920

RESULT 4
ADM40772
ID      ADM40772 standard; DNA; 1923 BP.
XX
AC      ADM40772;
XX
DT      01-JUL-2004 (first entry)
XX
DE      DNA encoding alkaline protease from Bacillus sp. KSM-KP43.
XX
KW      alkaline protease; laundry detergent; bleaching agent; detergent;
XX      denature-cleaning agent; ds; gene.
XX
OS      Bacillus sp.; KSM-KP43.
XX
FH      Key
XX      Location/Qualifiers
XX      CDS
XX      1..1923
XX      /tag= b
XX      /product= "Bacillus sp. KSM-KP43 alkaline protease"
XX      sig_peptide
XX      1..618
XX      /tag= a
XX      /product= "Alkaline protease prepro sequence"
XX      mat_peptide
XX      619..1923
XX      /tag= c
XX      /product= "Mature alkaline protease"
XX
PN      US2004072321-A1.
XX
PD      15-APR-2004.
XX
PF      09-JUN-2003; 2003US-00456479.
XX
PR      26-JUN-2002; 2002JP-00186387.
XX      18-OCT-2002; 2002JP-00304232.
XX
PA      (KAOS ) KAO CORP.
XX
PI      Sato T, Okuda M, Takimura Y, Sumitomo N, Nomura M, Kobayashi T;
XX

```


QY 562 SerThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
DB 1687 AGACAACTGCTCCGTAACGCTTGTCAATGATCTGACCTTGTCACTACCGCTCCAAAT 1746
QY 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnTrpAspGlyArg 601
DB 1747 GGCACACAGATATGTAGAAATGACTTACTTCGCCATACAAATGATTAATCGGATGCGCCG 1806
QY 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
DB 1807 AATAACGTGAAAAATATATTATTATTAATGACCAACAAAGCGGACGATACAAATTTGAGTA 1866
QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
DB 1867 CAGGCTTAATACCTACCGGTGACACACACCTTCTCGTGGCAATGTGAAT 1920
RESULT 5
ADSI4428
ID ADSI4428 standard; DNA; 1923 BP.
XX
XX ADSI4428;
AC
XX 30-DEC-2004 (first entry)
DT
XX Bacillus alkaline protease KP43 gene SEQ ID NO:2.
DE
XX protease; enzyme; alkaline protease; laundry detergent; KP43; ds; gene.
KM
XX Bacillus sp.; KSM-KP43.
OS
XX
XX Key Location/Qualifiers
FH 1.1923
FT CDS /tag= D
FT sig_peptide /product= "alkaline protease"
FT 1.618
FT mat_peptide /tag= a
FT 619..1920
FT /tag= C
XX
XX EPI466962-A1.
XX
XX 13-OCT-2004.
PD
XX
XX 08-APR-2004; 2004EP-00008604.
PF
XX 10-APR-2003; 2003JP-00106709.
PR
XX
XX (KAOS) KAO CORP.
PA
XX Okuda M, Sato T, Takimura Y, Sumitomo N, Kobayashi T;
PI WPI; 2004-711313/70.
DR P-PsDB; ADSI4437, ADSI4427.
XX
XX
XX New engineered alkaline protease, useful particularly in laundry
PT detergents, comprising specified amino acids at particular positions.
XX
XX
XX Disclosure; SEQ ID NO 2; 31pp; English.
PS
XX
XX The invention relates to a novel alkaline protease. The new alkaline
CC protease comprises an amino acid sequence in which one or more amino acid
CC residues selected from those located at 7 specific positions within
CC ADSI4427, or at positions corresponding to these positions are: position
CC 15 (histidine), position 16 (threonine or glutamine), position 166
CC (glycine), position 167 (valine), position 187 (serine), position 346
CC (arginine), and position 405 (aspartic acid). The alkaline protease is
CC useful in industry particularly in laundry detergents, but also e.g. in
CC fibre modifying agents, leather processing agents, cosmetic compositions,
CC bath additives, food-modifying agents, and pharmaceuticals. The present
CC sequence encodes the wild-type Bacillus sp. KSM-KP43 alkaline protease,
CC KP43.
XX
XX Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.35e-227 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 13 Gaps: 0
US-10-784-870-4 (1-639) x ADSI4428 (1-1923)
QY 2 ArgGlyValLeuValPheLeuSerValIleSerAlaAlaIleLeuSerThrValAla 21
DB 7 AAGAGAAAAAGGTGTTTATCTGTTTATCACTGCGACGCAATTTTGTGACTGTGGC 66
QY 22 LeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheGlyIleGln 41
DB 67 TTAAGTAAATCAATCTGACAGTGTGCAAGAAATTTGATCTGATTTCAAGAAATTCAG 126
QY 42 ThrThrThrAspValSerGlyPheSerGlyGlnArgGlnThrGlyAlaAlaIleLeu 61
DB 127 ACAACAACGATGCTAAAGGTTTCTCAAGCAGGCGCAGACTGTGCTGCTTTTCG 186
QY 62 ValGluSerGluAsnValIleLeuLeuLeuGlyLeuLeuValIleLeuValPro 81
DB 187 GTGAAATCTGAATAATGTGAACCTCCAAAGGTTTGCAAGAAAGCTTGAACAGTCCG 246
QY 82 AlaAsnAsnValLeuHisIleValGlnPheAsnGlyProIleLeuGluGluThrIleGln 101
DB 247 GCAAATTAATTAACCTCAATATTAATCAATTCATATGACCAATTTTGAAGAAACAAACG 306
QY 102 LysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
DB 307 CAGCTGAAAAAAGCGGGCAAGATTCGACTCATACCTGAATTAATGCTTAACATTCG 366
QY 122 GluTyrGluGlyAspValGluSerLysValArgSerIleGluHisValGluSerValGlu 141
DB 367 GAGTATGAGGCGGAGTTTAATGACACACACACATTAAGACGAGTGAATCCGTGAG 426
QY 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThrLeu 161
DB 427 CCTTATTTGCCGATTAACAGAAATAGATCCCGCTTTTCACAAAAGGGGCATCAGACTT 486
QY 162 ValIleValAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgGlyIle 181
DB 487 GTMAAAGCAGTCCCTGATCAAAAGCAAAATTAAGAGTCAATTAAGAGGCATC 546
QY 182 GluGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaLysProGlu 201
DB 547 GAACAATGCAACAATTCGCAATTAAGCAATGATGTGCTATATTAACGCAAAAGCCTGAG 606
QY 202 TyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSer 221
DB 607 TTTAAGGTGATGAATGATCTTCCGCTGGAATTCGAAAGCGGAGTGTGCTCAGAGCAGC 666
QY 222 TyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
DB 667 TACGGGTGTGATGACAAGACAGATCGTAGCGGTGCGATACAGAGGCTTGATCAGGT 726
QY 242 ArgAsnAspSerSerMetHisGluAlaPheArgGlyLysValIleThrAlaLeuTyrAlaLeu 261
DB 727 GCAATATGACAGTTCATGATCAATGAACCTTCCGCGGAAATTAATGCAATTAATATCATTTG 786
QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerVal 281
DB 787 GGAAGCAAGATTAATGCAATGATGCAATGATGCAATGATGCAATGATGCAATGATGCAAT 846
QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
DB 847 TTAGGAAAACGCTCCATTAATTAAGAAATGCGGCTCAGGCGCAATCTAGTCTCCAACT 906
QY 302 IleMetAspSerSerGlyGlyLeuGlyLysProSerAsnLeuGlnThrLeuPheSer 321

```

Db      907 ATCATGATAGCGGTGGGAGCTTGAGGAGCTACCTTCGAAATCTGCAACCTTATTCAGC 966
Qy      332 GlnAlaPheSerAlaGlyValAlaArgIleHsThrAsnSerTrpGlyValAlaIleValengly 341
Db      967 CAGGCAATACAGTGTGGTCCAGAAATTCATACAACTCTGCGAGGACGACATGAAATGCG 1026
Qy      342 AlaTyThrThraPheSerArgAsnValAlaPhePyrValAlaArgIleAsnAspMetThrIle 361
Db      1027 GCTTACACACAGATTCGAAATGCGATGACTATGTCGCGAAATATATATGACGATC 1086
Qy      362 LeuPheAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAla 381
Db      1087 CTTTTCGCTGCGCGGATGAAGACCGAAGCGCGAATCATGTCACAGGCAAGCT 1146
Qy      382 LysAsnAlaIleThrValGlyValAlaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
Db      1147 AAAAATGCAATAAAGTCCGAGCTACGAAAACTCCGCGCAAGCTTGGGTCTTATGGG 1206
Qy      402 AspAsnIleAsnHsValAlaGlnPheSerSerArgGlyProThrIleAspGlyArgIle 421
Db      1207 GACATATCAACCATGTGGCAAGTCTCTTCACGTCGACCGAACAAGATGACGATC 1266
Qy      422 LysPheAspValMetAlaProGlyTyrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
Db      1267 AAACCGAGATGTCATGGCAACGCGGAACGTTCAATACATCAGCAAGATCTTCTTGCAACG 1326
Qy      442 AspSerSerPheThrAlaAsnHsIleAspSerIleTyrAlaTyrMetGlyGlyThrSerMet 461
Db      1327 GATTCCTCTCTTGCGCGAACAATGACATGAATATGATTAATGATGAGTGAACGTCCTATG 1386
Qy      462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHsPheValIleAsnArg 481
Db      1387 GGTACACCGATGCTGTGGAACAGTCGACAGCTTCGTCGACATTTGTGAAAAACGA 1446
Qy      482 GlyIleThrProLysProSerLeuLeuValAlaLeuIleAlaGlyAlaAlaAspVal 501
Db      1447 GGCATCACACCAAGCTTCTCTAATTAAGCGGACATGATTCGCGTGACCTACATC 1506
Qy      502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrTrpGlyArgValThrLeuAspLysSerLeu 521
Db      1507 GGCCTGTGCTACCGCAACGATACCAAGATGGAGCGATGATTAATCCCTCG 1566
Qy      522 AsnValAlaTyValAsnGlnSerSerAlaLeuSerThrSerGlnIleValAlaThrTyrThr 541
Db      1567 AACGTGTGCTATGTAACAGATCCAGTCTCTATCCACAGCCAAAAAGCCAGTACTCG 1626
Qy      542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTrpSerAspAlaProAla 561
Db      1627 TTTACTGCTACTGCGCGCAAGCTTTGAATAATCTCCGTGATGCTGATGCCCTGCG 1686
Qy      562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db      1687 AGCAACAACGCTTCGCTAAACGCTGTCAATGATCGAACCTTGATTAACCCCTCAAT 1746
Qy      582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTrpAspGlyArg 601
Db      1747 GGCACACAGTATGTAAGAAATACTTACTTCGCCATACAAAGATACCTGGATGGCCGC 1806
Qy      602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyTyrThrTyrThrIleGluVal 621
Db      1807 AATAACGTATGAATAATGATATTATTAATGATGACCAACAAGCGGAGCTATATCAATTGAGCTA 1866
Qy      622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db      1867 CAGGCTATTAACGTAACGGTGGACCAAGACCTTCTCGTGGCAATGTGAAT 1920

```

```

XX      XX      Bacillus alkaline protease encoding DNA.
DE      DE      Alkaline protease; detergent; fiber modification; leather processing;
XX      XX      cosmetic; bath additives; food-modification; pharmaceutical; gene; ds.
KM      KM      Bacillus sp. KSM-KP43.
KW      KW      Key
XX      XX      CDS
XX      XX      Location/Qualifiers
XX      XX      1..1923
XX      XX      /tag= b
XX      XX      /product= "alkaline protease"
XX      XX      /note= "Seq ID 10"
XX      XX      sig_peptide
XX      XX      1..618
XX      XX      /tag= a
XX      XX      /note= "signal peptide"
XX      XX      /tag= c
XX      XX      /product= "mature alkaline protease"
XX      XX      /note= "Seq ID 1"
XX      XX      EPI466970-AL.
XX      XX      13-OCT-2004.
XX      XX      PD
XX      XX      08-APR-2004; 2004EP-00008605.
XX      XX      PF
XX      XX      10-APR-2003; 2003JP-00106708.
XX      XX      PR
XX      XX      (KAOS ) KAO CORP.
XX      XX      Sato T, Okuda M, Koyama S, Izawa Y, Kobayashi T;
XX      XX      WPI; 2004-711317/70.
XX      XX      DR      P-PSDB; ADT49604, ADT49613.
XX      XX      PT      New engineered alkaline protease with improved activity and thermal
XX      XX      stability, useful particularly in detergents such as laundry detergents.
XX      XX      Example 2; SEQ ID NO 2; 35pp; English.
XX      XX      PS      Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;
XX      XX      SQ
XX      XX      Alignment Scores:
XX      XX      Pred. No.: 1.35e-227 Length: 1923
XX      XX      Score: 3065.00 Matches: 595
XX      XX      Percent Similarity: 96.9% Conserved: 23
XX      XX      Best Local Similarity: 93.3% Mismatches: 20
XX      XX      Query Match: 93.8% Indels: 0
XX      XX      DB: 13 Gaps: 0
XX      XX      US-10-784-870-4 (1-639) x ADT49605 (1-1923)
Qy      2 ArgIleLysValPheLeuSerValLeuSerAlaAlaIleLeuSerThrValAla 21
Db      7 AAGAAAGAAAAAGGTGTTTATCTGTTTATCACTCCAGGATTTGTCACTGTGGC 66
Qy      22 LeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
Db      67 TTAAGTATTCATCTGCAAGGTGTCGCAAGAAATTTGATTCGATTCGAAAGAAATTCAG 126
Qy      42 ThrThrThrAspValSerGlyPheSerIleGlnArgGlnThrGlyAlaAlaAlaPheLeu 61
Db      127 ACAACAATGATGCTAAAGGTTTCTCAAGACAGGCGACGACTGCTGCTGCTTTCTG 186

```

QY	62	ValIGInuseGjUaenValIybeUeulv6GjYleuUeulv6YleuGjUthrValPro	81
Db	187	GTGGAAATCGAAAAATGTGAAATCTCCCAAAAGGTTTGCAGAAAGAGCTTGAAACAGTCCCG	246
QY	82	AlaAenAnLyLeuHsiIleValGInPheAnGjUProGleuGjUthrUyGIn	101
Db	247	GCAATATATAACTCCATATATACCAATTCAMATGAGCAATTTTAAAGAAACAAACAG	306
QY	102	IybeUeGjUthrUyAlaUySiIleUeApyrTleProApyrTAlaTyrIleVal	121
Db	307	CAGCTGGAAAAAACAGGGGCAAAAGATTCCTGACTACATACCTGATTATGCTTACATTC	366
QY	122	GjUyTyrGjUgJYAbpValGInserUyValArgSerTleGjUuhIevaGjUserValGjU	141
Db	367	GAGTATGAGGGGGAATGTTAAGTCAGCAACAGCAACCAATTCAGCACTGGAAATCCGTGAG	426
QY	142	ProTyrUeProUeTyrUySiIleApyProGjUnUePheThrUyGjUyAlaSerThrUeU	161
Db	427	CTTATTTGCCATATACGAATATAGATCCCAAGCTTTTACAAAAAGGGCAATCAGAGCTT	486
QY	162	ValIyValaIeUeAlaUeApyrThrUyGInAenAnLyGjUvAlGInUeApyrGjUyIle	181
Db	487	GTAAGAAGCACTGGCGCTTATACAAAGCAAGAAATTAAGAGCTGCATTAAGAGGCATC	546
QY	182	GjUgJuiIeAlGInUyValaIeSerAenApyValaHsTyrIleThrAlaUyProGjU	201
Db	547	GAACAAAATCGCAAAATTCGCAATTAAGCAATGATGTGCTATATATACCGCAAAAGCTGAG	606
QY	202	TyrUyValaIeAenApyValaIaArgGjUyIleValUyValaApyValaAgInSerSer	221
Db	607	TATAAGGTGATGAATGAATGTTCGGGTGAATTTGCMAAGCGAATGTGCTCAGAGCAAC	666
QY	222	TyrGjUyUeUyGjUyGInUyGInIleValaIaAvalaApyrThrGjUeUeApyrThrGjU	241
Db	667	TACGGCTTGTATGACACAGACAGATCGTAGCGGTTCGCAATACAGGGCTGTATACAGGT	726
QY	242	ArgAenApySerSerMechIeGjUuIaPheArgGjUySiIleThrAlaUeUyTyrAlaUeU	261
Db	727	CGCAATGACAGTTTCAGTACATGAAGCGCTTCGCGGGAATTAATCTGCAATTATATGATTTG	786
QY	262	GjUyAgtThrAenAnLyAaenApyrThrAenGjUyHAgjUyThiValaIaGjUySerVal	281
Db	787	GGAACGAAGAAATGAATGCCAATGATACGAATGTCTCATGTGACCAATGTGGCTGCTCCGTA	846
QY	282	LeuGjUyAenGjUyAlaThrAenUyGjUyMetAlaProGjUnAlaAenUeValaPheGInSer	301
Db	847	TTAGGAACCGGCTCCATATATAAAGAAATGGCGCTCAGGCGCAATCTAGTCTTCCAAATCT	906
QY	302	IleMeApySerSerGjUyGjUeUyGjUyUeUeProSerAenUeGjUthrUePheSer	321
Db	907	ATCATGATAGTAGGGCGGTGGGGACTTGGAGGACTACTTGCMAATCTGCMAACTTATTCACG	966
QY	322	GjUuAlaPheSerAlaGjUyAlaArgTleHsThrAenSerTyrGjUyAlaAlaValaAenGjU	341
Db	967	CAAGCATTCAGTGTGCTGTGCCAATTCATCAAACTCGGGGAGAGCACTGATGATGGG	1022
QY	342	AlaTyrThrThrApySerThrAenValaApyrTyrValaArgUyAaenApyrMechIle	361
Db	1027	GCTTACACAAACGATTTCCAGAAATGTGATGACTATGTGCGCAAAATATATATGACAGATC	1086
QY	362	LeuPheAlaAlaGjUyAenGjUuArgProAenGjUyTyrIleSerAlaProGjUyThrAla	381
Db	1087	CTTTTCGCTGCGGGAAATGAAGACCTGAACGGCGGAACCATAGTGTCAACAGGCAAGCTT	1144
QY	382	UyAaenAlaIleThrValaGjUyAlaThrGjUaenUeUyArgProSerPheGjUySerTyrAla	401
Db	1147	AAAAATGCAATTAACAGTGGAGCTACGAAAACTCCGCGCAAGCTTTGGGTCTTATGCG	1206
QY	402	AerAenIleAenIleValaIeGInPheSerSerArgGjUyProThrUyAerGjUyArgIle	421
Db	1207	GACAAATATCAACCAATGTGCAACAGTTCTTCAACGTGACCGCAAAAGATGAGAGGATCT	1266

QY	4422	LYSPROAPAPVAlmclAlAPROGlyThrTYrIleuSerAlaArgSerSerLeuAlaPro	441
Db	1267	AAACCGAGTGCATGGCAACGGGAAACGTTCACTATCAGCAAGATCTTCTTGGACCG	1326
QY	4442	AAPSerSerPheTPAlaAsnHlAASPSeLySTyAlaTYrMetGlyTYrSeMet	461
Db	1327	GATTCCTCTCTCTGGGCAACCATGACATGTAATATGCAATCATGGGTGGAACTCATG	1386
QY	462	AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHlApeValLysAsnArg	481
Db	1387	GCTACACCGATCGTGTGCTGGAAACGTGGCACAGCTTCGTGAGCATTTTGGAAAAACAG	1446
QY	482	GlyIleThrProLysProSerLeuLeuValAlaLeuIleAlaGlyAlaAlaPAPVal	501
Db	1447	GGCATTCACACCAAGCCCTTCTCTATTTAAAGCCGCACTGATGGCGGTGACGTCGATC	1506
QY	502	GlyLeuGlyTYrProAsnGlyAsnGlnGlyTropAlaArgValThrLeuAspLysSerLeu	521
Db	1507	GGCTTGGCTACCCGAAACGGTACCAAGATGGGACAGATGACATTGATTAATCCCTG	1566
QY	522	AsnValAlaTYrValaGlnJuseSerAlaLeuSerThrSerGlnLysAlaThrTYrThr	541
Db	1567	AACGTGCTATGTGAAAGAGTCCAGTTCCTATCCACCAAGCCAAAGAGGAGTACTCG	1626
QY	542	PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAla	561
Db	1627	TTTACTGCTACGCGCGCAAGCCCTTTGAAATCTCCCTGATACGTCTGATGCCCCGCG	1686
QY	562	SeThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn	581
Db	1667	AGCAACAACGCTTCGTAACGCTTGCAATGATCTGACCTTGTCATTACCGCTCCAAAT	1746
QY	582	GlyThrArgTYrValGlyAsnAspPheSerAlaProPheAspAsnAsnTPAAPGlyArg	601
Db	1747	GGCACACAGATGTAGAAATGACTTTACTTCGCCATACATATATATGGATGGCGCGC	1806
QY	602	AsnAsnValGlnAsnValPheIleAsnSerProGlnSerGlyThrTYrThrIleGluVal	621
Db	1807	AATAACGTAGAAATGTATTTATTAATGACACCAAGCGGAGCGTATCAATTGAGGTA	1866
QY	622	GlnAlaTYrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn	639
Db	1867	CAGGCTTATTAACGTACCGGTGGACCAAGACCTTCTCGTTGGCAATGTGAAT	1920
RESULT 7			
AAV82382			
ID	AAV82382 standard; DNA; 3003 BP.		
XX			
AC	AAV82382;		
XX			
DT	12-APR-1999 (first entry)		
XX			
DB	Bacillus JPI70 protease gene.		
XX			
KW	Protease; detergent; surfactant; leather processing; debittering;		
KW	flavour; ss.		
XX			
OS	Bacillus sp.		
XX			
XX			
FN	Key		
FT	Location/Qualifiers		
FT	CDS		
FT	846..2771		
FT	/*tag= a		
FT	sig_peptide		
FT	846..944		
FT	/*tag= b		
FT	mat_peptide		
FT	1470..2768		
FT	/*tag= c		
XX			
XX			
XX	MO9856927-A2.		
XX			
XX	17-DEC-1998.		
XX			
XX	09-JUN-1998; 98WO-US012005.		
XX			

PR 12-JUN-1997; 97US-00873479.
XX (NOVO) NOVO NORDISK BIOTECH INC.
XX Sloma A, Christiamson L;
XX MPI, 1999-080908/07.
DR P-PSDB; AAM89547.
XX Novel protease from Bacillus subtilis LC20 - useful in laundry and
PT dishwashing detergents and for leather processing.
XX
XX Claim 11; Page 52-53; 77pp; English.
XX This nucleotide sequence encodes a novel protease (see AAM82382) of
CC Bacillus sp. JP170 (NCIB 12513). The sequence in plasmid p170BAN is
CC contained in Bacillus subtilis LC20 NRRL B-21680. The protease gene was
CC isolated from chromosomal DNA of JP170 following preparation of probes
CC based on protease N-terminal and internal peptides (see AAM89549-50),
CC screening of chromosomal libraries, isolation of the 3' end of the gene
CC by inverse PCR (see AAM82410-11), reconstruction of 5' and 3' ends and
CC PCR amplification (see AAM82412-16). Claimed recombinant host cells can
CC be used in a method for producing the protease. The protease is used in
CC laundry and dishwashing detergents, for institutional and industrial
CC cleaning, and for leather processing, as well as for debittering and
CC enhancing the degree of hydrolysis of protein hydrolysates, for flavour
CC development through hydrolysis of proteins, degradation of undesired
CC peptides and in enzymatic synthesis of peptides. It has enhanced
CC stability towards oxidation under alkaline conditions, e.g. towards
CC bleaching agents of the peroxy type. The invention also provides mutant
CC cells in which the protease activity is diminished. Such cells can be
CC used for the production of heterologous recombinant proteins
XX
SQ Sequence 3003 BP; 982 A; 504 C; 645 G; 872 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,09e-206 Length: 3003
Score: 2799.00 Matches: 541
Percent Similarity: 92.4% Conservatave: 52
Best Local Similarity: 84.3% Mismatches: 45
Query Match: 85.7% Indels: 4
DB: \. Gaps: 2
US-10-784-870-4 (1-639) x AAM82382 (1-3003)
QY 1 MecArgLys-----LysLysValPheLeuSerValLeuSerAlaAlaIleLeu 17
Db ATGAGAAAGAAAGATCGAAGAGGGCTTTTATCCGTTTATCAGTTCGCACTATTG 905
QY 18 SerThrValAlaLeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPhe 37
Db TCTTCTGTGCTTAAAGCATCTCTTACTATTGGGGGAAACATTTGAAATGGACTTT 965
QY 38 LysGlyIleGlnThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAla 57
Db AAGGGGATAGACACCTTACGCTTAGAAGAGCTGCCACCAAGCAAGAAAAAGGGAAG 1025
QY 966 AAGGGGATAGACACCTTACGCTTAGAAGAGCTGCCACCAAGCAAGAAAAAGGGAAG 1025
Db 58 AlaAlaPheLeuValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLeu 77
QY 1026 GCATCTTTCTTGTAACTCTGAAATATGTAATATCCCAAGAGATTCAAAGAACTA 1065
Db 78 GluThrValProAlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlu 97
QY 1086 GAAGTAGTTCCAGCGGATTAACAAGCTATATACGTTCAATTTGAGAGCACTAATTTTAAAG 1145
Db 98 GluThrLysGlnLysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyr 117
QY 1146 GAAGCGCACTTCAACTAGAGAAAGCGGAGCGAAATTTCTGATTACCTACCAATTAC 1205
Db 118 AlaTyrIleValGluTyrGluGluLysAspValGlnSerLysValArgSerIleGluHisVal 137
QY 1206 GCTTATATGTCGAAATATGATGGGATGTAAGAGCGCTAATCAAGCAATTTGCGCATTTTG 1265

QY 138 GluSerValGluProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGly 157
Db 1266 GAATCGGTGAAACCAATTTTACCTTTATATTAATTAACCCGCAATTTTCCAGAGCA 1325
QY 158 AlaSerThrLeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGlyValGln 177
Db 1326 GCTTCTGAATTAGAGAAACAAGTACGCTTTAGATAAAGCAAGAGATAAAGAGTACGT 1385
QY 178 LeuArgGlyIleGluGluIleAlaGlnTyrValAlaSerAspAspValHisTyrIleThr 197
Db 1386 TTAAGAGATTAAGAAATATGCGCAATTAAGCCGCAAAATATATGTTATATAGTACC 1445
QY 198 AlaLysProGluTyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspVal 217
Db 1446 CCAAGCCTGAATTCGAAGTTTAAATGACGTGGCCCGTGCATTGGAAGCAGACGTC 1505
QY 218 AlaGlnSerSerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaAlaAspThrGly 237
Db 1506 GCAAAATAATACCTTGGCTTATATGACAAAGACAGATTTGACATTGCTGATACGCGG 1565
QY 238 LeuAspThrGlyArgAsnAspSerSerMetHisGluLysPheArgGlyLysIleThrAla 257
Db 1566 CTGATACAGGAAAGAAATGACAGTTCCATGATGATGACATTTCCGCGTAAAGATTACCGCA 1625
QY 258 LeuTyrAlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisVal 277
Db 1626 CTAATATGCACTGGGCAAGACAAATACAGCCAAATATCCAAATGACATGAAACCAATCTT 1665
QY 278 AlaGlySerValLeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeu 297
Db 1686 GCTGATCTGTGTTAGAAAT--GTCACAAATTAAGAGATGGCAGCCGCAAGCAATCTA 1742
QY 298 ValPheGlnSerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGln 317
Db 1743 GTCTTCAATCTAATTATGATATGATGATGAGGCGCTGAGAGACTTACCTGTAATCTCA 1802
QY 318 ThrLeuPheSerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTrpGlyAla 337
Db 1803 ACATTAATTCAGTCAAGCAATATATGCTGAGACGAAATTCATACAAATTCATGGGGGCT 1862
QY 338 AlaValAsnGlyAlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsn 357
Db 1863 CCAATAAACGGTGCCTATACAGACAGCTCTCGAAATGTTGATGATTATGCAAGAAAT 1922
QY 358 AspMetThrIleLeuPheAlaAlaGlyAsnGlyLysArgProAsnGlyTyrThrIleSerAla 377
Db 1923 GATTTAGCAATCTTTTGGCGCGGAAATGAGGAGCCAGTACCGGTACATACAGTCA 1982
QY 378 ProGlyThrAlaLysAsnAlaIleThrValGlyValaThrGluAsnLeuArgProSerPhe 397
Db 1983 CCAAGAACAGCAAAATATGCAATTAACGTTTGGGCAACCAAAATCTACGTCAGACCTTC 2042
QY 398 GlySerTyrAlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLys 417
Db 2043 GATCTTATGGGATATATTAATTAACATGTTGCTCAATTCCTTCAAGAGGCTCACTAGA 2102
QY 418 AspGlyArgIleLysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSer 437
Db 2103 GATGAGAGTATTAAGCCGAGCGTCAATGCAACAGTACGATATATCTCTGCTAGATCA 2162
QY 438 SerLeuAlaProAspSerSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMetGly 457
Db 2163 TCATTACCTCCAGATTCCTCAATTTGGGCAAAACATATATATATATATGCTTACATGGGT 2222
QY 458 GlyThrSerMetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPhe 477
Db 2223 GGTACTTCTATGCTACTCAATTTGTAAGAGTAAATTTTGAACAATTAAAGGAGCATTTT 2282
QY 478 ValLysAsnArgGlyIleThrProLysProSerLeuLeuLysAlaAlaLeuIleAlaGly 497
Db 2283 GTGAAATATAGAGGGGTAACCTCAACCTTCCCTTTAAAGCTCTTAAATATGCAAGGT 2342
QY 498 AlaAlaAspValGlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeu 517

```
DB 2343 GCGGGGAGTGTGACTTGCTTCAGATGCGAACGAGTGGGAGGATACGTTA 2402
QY 518 AaplySerLeuAenValAlaTyrrValAenGluSerSerAlaLeuSerThrsGlnTyr 537
DB 2403 GATTAATCCCTTAATGCTTCCATTTGTAATGAAAGCCCTTTATCAACAAGTCAAAA 2462
QY 538 AlaThrTyrrThrsPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSer 557
DB 2463 GCAACATATTCGTTTACGGCTCAAGCTGTAAACCTTTAAATATCATCTTGTGGTCA 2522
QY 558 AapAlaProAlaSerThrsThrsAlaSerValThrLeuValAapAapLeuAapLeuValIle 577
DB 2523 GATGACACCGTAGACACGACGCGATCATCACTTATGTAATGATTTAAGCTTAAGTAAC 2582
QY 578 ThrAlaProAenGlyThrsArgTyrrValAlaGlyAapAapPheSerAlaProPheAapAapAen 597
DB 2583 ACTGCACCAATAGAACTTAATACGTGCGAAATGACTTTACAGACCGTATGATACAT 2642
QY 598 TTPAAPGlyArgAapAapAenValAlaGlyAenValPheIleAenSerProGlnSerGlyThrsTyr 617
DB 2643 TGGGATGGGAGAAACACGTGAAATGTTTATCAATGCTCTCAAGCGGAACGTAT 2702
QY 618 ThrIleGlyValAlaIleAlaTyrrAenValProValAlaGlyProGlnAapPheSerLeuAlaIle 637
DB 2703 ACGAGTCGAAGTCAGGCTTACATGATGACAGTAAGTCGCAAACTTTTCTTACGCAAT 2762
QY 638 ValAen 639
DB 2763 GTACAT 2768
```

RESULT 8

```
ADY33779
ID ADY33779 standard; DNA; 1305 BP.
```

```
XX AC ADY33779;
XX 05-MAY-2005 (first entry)
XX DT
XX DE Bacillus species alkaline protease coding sequence.
XX KM mutagenesis; alkaline protease; detergent; laundry; cosmetic; food;
XX KM bleaching detergent; gene; ds.
XX OS Bacillus sp.
XX
XX Key Location/Qualifiers
XX CDS 1..1305
XX FT /*tag= a
XX FT /product= "alkaline protease"
XX FT /note= "no start codon given"
XX FT /partial
XX
XX BP1347044-A2.
XX
XX 24-SEP-2003.
XX
XX 21-MAR-2003; 2003BP-00006472.
XX
XX 22-MAR-2002; 2002JP-00081428.
XX PR 06-JUN-2002; 2002JP-00155987.
XX PR 18-OCT-2002; 2002JP-00304230.
XX PR 18-OCT-2002; 2002JP-00304231.
XX
XX PA (KAOS ) KAO CORP.
XX
XX Ohnda M, Sato T, Saito K, Sumitomo N, Izawa Y, Saeki K;
XX Kobayashi T, Nomura M;
XX WPI; 2003-846540/79.
XX DR P-PSDB; ADY33778.
XX
XX New alkaline protease having specific amino acid residue at a specific
```

PT position of its amino acid sequence, useful for producing detergent
PT compositions, laundry detergent, fiber modifiers, leather-treating agents
PT or pipe cleaners.

Disclousre; SEQ ID NO 2; 31pp; English.

CC The invention relates to an alkaline protease having a fully defined
CC sequence of 434 amino acids (I) given in the specification, or an amino
CC acid sequence at least 80% homology with (I), where an amino acid residue
CC at position 65, 101, 163, 170, 171, 273, 320, 359 or 387 of (I) is
CC selected from 16 amino acid residues. The amino acid residues at the
CC corresponding positions are selected from: position 65: proline; position
CC 101: asparagines; position 163: histidine; aspartic acid, phenylalanine,
CC lysine, asparagine, serine, isoleucine, leucine, glutamine, threonine or
CC valine; position 170: valine or leucine; position 171: alanine, glutamic
CC acid, glycine or threonine; position 273: isoleucine, glycine or
CC threonine; position 320: phenylalanine, valine, threonine, leucine,
CC isoleucine or glycine; position 359: serine, leucine, valine, isoleucine
CC or glutamic acid; and position 387: alanine, lysine, glutamine, glutamic
CC acid, arginine or histidine. The alkaline protease is useful for the
CC production of a detergent composition, such as laundry detergent, fiber
CC modifiers, leather-treating agents, cosmetic compositions, bath
CC additives, food modifiers and pharmaceutical compositions. The alkaline
CC protease may also be used as bleaching detergent, hard surface cleaning
CC detergent, pipe cleaner, artificial tooth cleaner, and as a sterilizing
CC cleanser for medical tools. The new alkaline protease has a more potent
CC proteolytic capacity, exhibiting excellent detergency for the removal of
CC a complex stain, and has high secretion capacity. This sequence
CC corresponds to the Bacillus sp. alkaline protease coding sequence.

SO Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;

Alignment Scores:

```
Pred. No.: 1,81e-159 Length: 1305
Score: 2183.00 Matches: 418
Percent Similarity: 99.34 Conservative: 13
Best Local Similarity: 96.34 Mismatches: 3
Query Match: 66.84 Indels: 0
Dn: 11 Gaps: 0
```

US-10-784-870-4 (1-639) x ADY33779 (1-1305)

```
QY 206 AapAapValAlaAargGlyIleValIleAapValAlaGlnSerSerTyrrGlyLeuTyr 225
DB 1 AATGATGTTGGCGGTGAATGTCMAAGCGATGTCGACAGACGCTACCGGTGAT 60
QY 226 GlyGlnGlyGlnIleValAlaValAlaAapThrsGlyLeuAapThrsGlyArgAapAapSer 245
DB 61 GGCACAGACAGATGTTAGCGGTTGCCATACAGGAGCTTGATACAGTGCATACACT 120
QY 246 SerMetHisGluAlaPheArgGlyLysIleThrAlaLeuTyrrAlaLeuGlyArgThrsAen 265
DB 121 TCGATGCATGAAGCTTCGCGGGAATTTCTGCAATTATATGATTTGAGACGACGAAAT 180
QY 266 AapAlaAapAapThrsAenGlyHisGlyThrsIleValAlaGlySerValLeuGlyAenGly 285
DB 181 AATGCCAATGATACGAATGTCATGTCATGTCATGTCGTCGCTCGTATTAGAAACGCG 240
QY 286 AlaThrAenLysGlyMetLeuAlaProGlnAlaAenLeuValPheGlnSerIleMetAapSer 305
DB 241 TCACATATTAATTAAGAAATGCGCTCAGCGCAATCTAGCTTCATCATCAATGATAGC 300
QY 306 SerGlyGlyLeuGlyGlyLeuProSerAenLeuGlnThrLeuPheSerGlnAlaPheSer 325
DB 301 GGTGGGGGACTTGGAGGACTACCTTCGAAATCTGCAAACTTATTCAGCCAAACATACAGT 360
QY 326 AlaGlyAlaAargIleHisThrsAenSerTyrrGlyAlaAlaValAenGlyAlaTyrrThrs 345
DB 361 GCTGTGTCAGAAATTCATACAACTCCCTGGGAGACGACGATGAAATGGGCTTACCAACA 420
QY 346 AapSerArgAenValAapAapTyrrValArgTyrrAenAapMetThrIleLeuPheAlaAla 365
DB 421 GATTCCAGAAATGTCAGATGATATGTCGCAAAATGATATGACATCTTTTTCGCTGCC 480
```



```
QY 366 GlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaIleAsnAlaIle 385
DB 481 GGGATGAGAGCCGACCGGAGACATCATGACACGACGACGCTAAATGCAATA 540
QY 386 ThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsn 405
DB 541 ACAATCGAGCTGACGGAACACTCGCCCAAGCTTTGGGTCTTATGCGGACAAATATCAAC 600
QY 406 HisValAlaGlnPheSerSerThrGlyProThrIleAspGlyAlaGlyIleArgProAspVal 425
DB 601 CATGGGACAGTCTCTTACAGTGGACCGACAAAGATGACGATCAACCGATATC 660
QY 426 MetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerPhe 445
DB 661 ATGGACCGGAGACCTTCACTACATCAAGATCTTCTTTCACCGGATCTCTCTTC 720
QY 446 TrpAlaAsnHisAspSerLeuTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIle 465
DB 721 TGGGCGAACCATGACATTAATATGATCATGATGGTGAACGTTCATGCGTACACGATC 780
QY 466 ValAlaGlyAsnValAlaGlnLeuArgGluHisPheValIleAsnArgGlyIleThrPro 485
DB 781 GTTGCTGGAAACGTGGACAGCTTCGTAGCATTTTGGAAAAACAGGACATCACACA 840
QY 486 LysProSerLeuLeuValAlaLeuIleAlaGlyAlaAlaAspValGlyLeuGlyTyr 505
DB 841 AAGCTTCTCTATTAAGGCGGACCTGATTCGCGGTGACGCTGACATGCGCTTGGCTAC 900
QY 506 ProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeuAsnValAlaTyr 525
DB 901 CCGAACCGTAAACCAAGATGGGAGAGATGACATGATTAATCCCTGAACGTTGCCAT 960
QY 526 ValAsnGlySerSerAlaLeuSerThrSerGlnValAlaThrTyrThrPheThrAlaThr 545
DB 961 GTGAACGAGTCAGTCTCTATCCACCAAGCAAAAAGGAGATGACTCTGTTTACTGCTACT 1020
QY 546 AlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAlaSerThrAla 565
DB 1021 GCCGCGACGCTTTGAAATCTCCCTGGATGGTGTATGATGCCCTTGCAGACATCTCT 1080
QY 566 SerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyThrArgTyr 585
DB 1081 TCCGTAAGCTTGTATGATGATCTGGACCTTTCATTAACGCTCCAAATGGCACACAGTAT 1140
QY 586 ValGlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGlyArgAsnAsnValGlu 605
DB 1141 GTAGGAATGACTTATCTTCCCATACATGATGATGCGATGCGCCGCAATPACGTAGAA 1200
QY 606 AsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluValAlaTyrAsn 625
DB 1201 AATGATATTATTAATGACCAACAAAGCGGAGATGATGAGTACAGGCTTATTAAC 1260
QY 626 ValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
DB 1261 GTACCGGTTGGACCAAGACCTTCTGTTGGAAATGTGAAAT 1302

RESULT 9
ADL25803
ID ADL25803 standard; DNA; 1305 BP.
XX ADL25803; .
XX
DT 20-MAY-2004 (first entry)
XX
DB Bacillus alkaline protease coding sequence.
XX
KW alkaline protease; washing agent; gene; ds; enzyme.
XX
OS Bacillus sp.
XX
FH Key
CDS Location/Qualifiers
FT 1..1305
```

```
FT /*tag= a
FT /partial
FT /product= "Bacillus alkaline protease"
FT /note= "No start codon is given"

JP2004008085-A.
XX
XX
XX
PD 15-JAN-2004.
XX
XX
XX
PP 06-JUN-2002; 2002JP-00165950.
XX
XX
XX
PR 06-JUN-2002; 2002JP-00165950.
XX
XX
XX
PA (KAO ) KAO CORP.
XX
XX
XX
DR WPI; 2004-094297/10.
XX
XX
XX
PT Novel mutant alkaline protease produced by substituting the amino acid
PT residues useful as washing agent.
XX
XX
XX
PS Disclosure; SEQ ID NO 2; 21pp; Japanese.
XX
XX
XX
CC The invention comprises a mutant Bacillus alkaline protease which
CC contains substitutions at positions 163, 170 and 434. The mutant alkaline
CC protease is useful as a washing agent. The present DNA sequence encodes a
CC Bacillus alkaline protease of the invention.
XX
SQ Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,81e-159 Length: 1305
Score: 2183.00 Matches: 418
Percent Similarity: 99.3% Conservative: 13
Best Local Similarity: 96.3% Mismatches: 3
Query Match: 66.8% Indels: 0
DB: Gaps: 0

US-10-784-870-4 (1-639) x ADL25803 (1-1305)

QY 206 AsnAspValAlaArgGlyIleValAlaAspValAlaGlnSerSerTyrGlyLeuTyr 225
DB 1 AATGATGTGGCGCTGGAATGTCAAGCGGATGTGCTCAGACGACGCTAGGTGTAT 60
QY 226 GlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSer 245
DB 61 GGAACAGACAGATCGTAGCGGTGGCCGATACAGGCTTGAATACAGGTCGCAATGACAGT 120
QY 246 SerMetHisGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeuGlyArgThrAsn 265
DB 121 TCGATGCAATGAAGCTTCCCGGGAATTAATGCAATTAATGCAATGGACGACCAAT 180
QY 266 AsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGlyAsnGly 285
DB 181 AATGCCATGATATGAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 286 AlaThrAsnIleGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSer 305
DB 241 TCCACTATATAAGGAAGAGGCGCTCAAGCGCAATCTTCAATCTTATCATGATGATGAT 300
QY 306 SerGlyIleLeuGlyGlyLeuProSerAsnLeuGlnIleThrLeuPheSerGlnAlaPheSer 325
DB 301 GGTGGGGGACTTGGAGGACATCACTTCCATCTGCAATCTTATTCACGCAAGCATACAGT 360
QY 326 AlaGlyAlaArgIleHisThrAsnSerTyrGlyValAlaValAsnGlyAlaTyrThrThr 345
DB 361 GCTGTGCCAGAAATTCATACAACTCTGGGGAGACAGATGATGGGCTTACACACA 420
QY 346 AspSerArgAsnValAlaAspTyrValArgLysAsnAspMetThrIleLeuPheAlaAla 365
DB 421 GATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 366 GlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaIleAsnAlaIle 385
```



```

DB 481 GGGAGTGAAGAGCCGGAACGGGAAACATCAATGCAACGACGACACTTAAATGCAAT 540
QY 386 ThrValGlyAlaThrGluLeuLeuArgProSerPheGlySerTrpAlaAspAsnIleAsn 405
DB 541 ACAAGTCGAGCTACGGAACAACTCCGCCAAGCTTGGGCTTATGCGGCAATATCAAC 600
QY 406 HisValAlaGlnPheSerSerArgGlyProThrIleValArgGlyArgIleLeuProAspVal 425
DB 601 CATGTGGCAAGTTCCTTCACGTGGACCGAACAAGAAATGACAGGATCAACACGGATGTC 660
QY 426 MetAlaProGlyThrTrpIleLeuSerAlaArgSerSerLeuAlaProAspSerPhe 445
DB 661 ATGGCAACCGGAAAGCTTCACTATACGAACATCTTCTTCGACCGGATTCCTCTTC 720
QY 446 TrpAlaAsnHisAspSerIleValArgIleValArgIleValArgIleValArgIleVal 465
DB 721 TGGGCGAACAATGACATGATATGATATGATATGATATGATATGATATGATATGATATG 780
QY 466 ValAlaGlyValAsnValAlaGlnLeuArgGluHisPheValLeuAsnArgGlyIleThrPro 485
DB 781 GTTGTGGGAAAGCTGACAGCTTCGTGAGCAATTTGTGAAAAACGAGGCAATCACACCA 840
QY 486 LysProSerLeuLeuValAlaLeuIleAlaGlyAlaAlaAspValGlyLeuGlyTrp 505
DB 841 AAGCTTCTCTATTAATAACCGGCACTGATTCGGGTGCGAGCTGACATCGGCTTGGCTAC 900
QY 506 ProAsnGlyAsnGlnGlyTrpGlyArgValIleLeuAspLysSerLeuAsnValAlaTrp 525
DB 901 CCGAAGCGGTACCAAGATGAGGGAAGATGACATTCGATTAATCCCTGAACGTTGCTAT 960
QY 526 ValAsnGluSerSerAlaLeuSerThrSerGlnValAlaThrTrpThrAlaThr 545
DB 961 GTGAACGAGTCCAGTTCTCTATCCACACGCAAAAACGACATCTGTTACTGCTACT 1020
QY 546 AlaGlyLysProLeuLysIleSerLeuValIleProSerAspAlaProIleSerThrAla 565
DB 1021 GCCGCAAGCTTTGAAATCTCCGTGATGCTCTGATGATGATGATGATGATGATGAT 1080
QY 566 SerValIleLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyThrArgTrp 585
DB 1081 TCCGTAACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 586 ValGlyAsnAspPheSerAlaProPheAspAsnAsnTrpAspGlyArgAsnAsnValGlu 605
DB 1141 GTAGAAATGACTTACTTCGCCATACATGATGATGATGATGATGATGATGATGATGAT 1200
QY 606 AsnValIleAsnSerProGlnSerGlyThrTrpIleGluValAlaIleValAsn 625
DB 1201 AATGATTTTATTAATGACCAACAAAGCGGAGCATTAATGAGGTACAGGCTTAAAC 1260
QY 626 ValProValGlyProGlnAsnAspSerLeuAlaIleValAsn 639
DB 1261 GTACCGGTGGACCAAGACCTTCTCGTTGGCAATTTGTAAT 1302

```

RESULT 10

```

AA027516
ID AA027516 standard; DNA; 1299 BP.
AC AA027516;
DT 05-FEB-1993 (first entry)
DE Alkali-protease Ya enzyme gene.
XX Alkali resistance; surface active agent resistance; detergency improver;
XX 88.
OS Bacillus sp. Y.
XX Key Location/Qualifiers
XX PH 1..1299
XX CDS
XX FT /*tag= a

```

```

XX JP04197182-A.
XX 16-JUL-1992.
XX 28-NOV-1990; 90JP-00327110.
XX 28-NOV-1990; 90JP-00327110.
XX (LION ) LION CORP.
XX
XX MPI; 1992-288440/35.
XX P-PSDB; AAR26274.
XX
XX DNA coding alkali-protease Ya enzyme - has good alkali and surfactant
XX resistance and improves detergency.
XX
XX Claim 3; Page 2; 17p; Japanese.
XX
XX The sequence is that of the alkali-protease Ya enzyme gene which can be
XX used in the recombinant production of Ya enzyme. Ya enzyme is excellent
XX in alkali resistance and surface active agent resistance and improves
XX detergency
XX
XX Sequence 1299 BP; 401 A; 246 C; 297 G; 355 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Sred. No.: 1,426-144 Length: 1299
XX Score: 1990.50 Matches: 381
XX Percent Similarity: 94.0% Conservative: 27
XX Best Local Similarity: 87.8% Mismatches: 25
XX Query Match: 60.9% Indels: 1
XX DB: 2 Gaps: 1
XX
XX US-10-784-870-4 (1-639) x AA027516 (1-1299)
QY 206 AsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSerTrpGlyLeuTrp 225
DB 1 AATGATGTAAGCAAGGAGGATGTAATAAGCTGATGTTGCACAAACATTAACGATTAAT 60
QY 226 GlyGlnGlyGlnIleValAlaValAlaAspTrpGlyLeuAspTrpGlyArgAsnAspSer 245
DB 61 GGCACAGGTCAACTGTTGCACTAGCGACGACACAGGCTTGAATACAGGTCGTAACATAT 120
QY 246 SerMetHisGluAlaPheArgGlyLysIleThrAlaLeuTrpAlaLeuGlyArgThrAsn 265
DB 121 TCTATGCAATGAAGCATTCGCGGGAATAATCAAGCTCTTACGCTTGAAGAACTAAT 180
QY 266 AsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGlyAsnGly 285
DB 181 AATGCAAGTGAATCCAAATGGGCATGCAACATGTAAGCAGTTCTGTATCTTGTAAT--- 237
QY 286 AlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValIlePheGlnSerIleMetAspSer 305
DB 238 GCTTTAAATTAAGAAATGGCTCCGCAAGCTAATTAAGCTTCAATCTAATTAATGAATAC 297
QY 306 SerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSer 325
DB 298 AGCGGAGGATTAAGTGGGCTTACCATCGAATTAATAATGATTAATGATTAATGATTAAT 357
QY 326 AlaGlyAlaArgIleHisThrAsnSerTrpGlyAlaAlaValAsnGlyValAlaTrpThr 345
DB 358 GCTGAGCAAGAAATTCATTAATCTTGGGAGGCCCAAGTAATGAGCGGACACTGCT 417
QY 346 AspSerArgAsnValAspAspTrpValArgLysAsnAspMetThrIleLeuPheAlaIle 365
DB 418 AACTGAAACAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 477
QY 366 GlyAsnGluArgProAsnGlyGlyTrpIleSerAlaProGlyThrAlaLysAsnAlaIle 385
DB 478 GGTATATGAAGGCTTAATTCAGAAACATTAATGATGATGATGATGATGATGATGATGAT 537
QY 386 ThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTrpAlaAspAsnIleAsn 405

```

```
Db 538 ACCGTCGGGCAACGGAAAATCTATGCCCACTTCGTTCCATATGACGATATACCCAAAT 597
Qy 406 HlsvAlaIaGlnPheSerSerArgIyProThrIyAaSpGlyArgIleIyProAspVal 425
Db 598 CATATTGCAATTTTTCATCGAAGAGCTACGAGGATGAGCAATTAAGCTTACGTA 657
Qy 426 MetAlaProGlyThrTyrlleuSerIlaArgSerSerLeuAlaProAspSerPhe 445
Db 658 ACAGCTTCCTGGAACTTTATTTATTCAGACGTTCTCTTCTTACCTCCAGACTTCCTGTT 717
Qy 446 TrpAlaAsnHisAspSerIySerIyAlaTyMetGlyIyThrSerMetAlaThrProIle 465
Db 718 TGGGGCAATTATACAGTAATACCGGTAATCGGGGTAATCTCCATGGCGACACTATT 777
Qy 466 ValAlaGlyAaenValAlaGlnLeuArgIyIhIAspValIyAaAspArgIyIleThrPro 485
Db 778 GTTGCAGGGAATGTGGCCGAATTACGTGAGCATTTTATTAATAAGAGATTAATCTCT 837
Qy 486 LysProSerLeuLeuIyAaIaLeuIleAlaGlyAlaAlaAspValGlyLeuGlyIy 505
Db 838 AAGCCTCTTTATTAATAAGCTCACTATCGCTGGTGAATGATGTTAGATAT 897
Qy 506 ProAsnGlyAaenGlyIyTrpGlyArgValThrLeuAspIySerLeuAaenValAlaTy 525
Db 898 CTTAGTGGTGACCAAGGCTGGGGGTGTACTCTAGAAATATCGTTAATGTACGTAAT 957
Qy 526 ValAaenGlySerSerAlaLeuSerThrSerGlnIyAlaThrTyrlleThrAlaThr 545
Db 958 GTCAATGAGCAACGTGCTTACCCACAGCAAAAAAGCAACGATATCGTCCACAGCA 1017
Qy 546 AlaGlyIyProLeuIyIleSerLeuValIyTrpSerAaPAlaProAlaSerThrAla 565
Db 1018 GCGGTTAACTTTAAATCTCGTTAGTATGAGCAAGTCTCGAAAGTACCACTGA 1077
Qy 566 SerValIyThrLeuValAaAspLeuAaPLeuValIleThrAlaProAsnGlyIyThrArgTy 585
Db 1078 TCTTATACACTAGTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1137
Qy 586 ValGlyAaenAspPheSerAlaProPheAaPAsnAaenTrpAspGlyIyArgAaenValGlu 605
Db 1138 GTAGAAATGATTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1197
Qy 606 AaenValIyPheIleAaenSerProGlnSerGlyIyThrTyrlleGlnValGlnAlaTyAsn 625
Db 1198 AACGATTTTAAAGCTCCGCAATCTGAGACGTATATTAATTGAGTTCAAGCGTAAAT 1257
Qy 626 ValProValGlyProGlnAaAspPheSerLeuAlaIleValAaen 639
Db 1258 GTACCATCTGGGCCACAGCGTTTCTCACTAGCTATCGTAAAT 1299

RESULT 11
AAT85667
ID AAT85667 standard; DNA, 1977 BP.
AC AAT85667;
XX
XX 17-OCT-2003 (revised)
DT 20-APR-1998 (first entry)
XX
XX Thermococcus protease coding sequence.
DE Thermococcus protease coding sequence.
XX
XX Protease; reasearch reagent; thermal stability; thermococcus celer; ss.
OS Thermococcus celer; DSM-2476.
XX
XX MO9721823-A1.
XX
XX 19-JUN-1997.
PD
XX
XX 07-NOV-1996; 96WO-JP003253.
XX
XX 12-DEC-1995; 95JP-00323285.
```

```
XX
PA (TAKI ) TAKARA SHUZO CO LTD.
XX
XX Takakura H, Morishita M, Yamamoto K, Mitra M, Asada K;
PI Tsunawawa S, Kato I;
XX
XX MPI; 1997-332794/30.
DR
DR P-PSDB; AAM24121.
XX
XX Protease(s) and genes encoding them obtained from Thermococcus and
PT Pyrococcus strains - have extremely high thermal stability and are useful
PR industrially and as research reagents.
XX
PS Claim 3; Page 86-87; 159pp; Japanese.
XX
CC This sequence represents the coding sequence for the protease from
CC Thermococcus celer DSM-2476. This sequence encodes a protease of the
CC invention. The proteases of the invention have extremely high thermal
CC stability. The proteases can be used as research reagents, and
CC industrially in the food, drug and chemical industries. (Updated on 17-
CC Oct-2003 to standardise OS field)
XX
SQ Sequence 1977 BP; 453 A; 659 C; 554 G; 311 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3,32e-27 Length: 1977
Score: 477.00 Matches: 168
Percent Simlarity: 41.8% Conservative: 90
Best Local Simlarity: 27.2% Mismatch: 203
Query Match: 14.6% Indels: 156
DB: 2 Gaps: 25
US-10-784-870-4 (1-639) x AAT85667 (1-1977)
Qy 62 ValGlnSerGlnAaenValIyIleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 80
Db 103 GTTCAGCAAGAACTACGAGCTGCTGACCCCGGAGCTGTTCAAGAAATGACAGAGATG 162
Qy 81 ProAlaAaenAaenIyIleuHisIleValGlnPheAaenGlyProIleLeuGlnGluThrIys 100
Db 163 AACTGAAACCGAAGAGGAGCAACCGCTAATATGTTCCGG----- 201
Qy 101 GlnIyIleuGlnThrThrGlnAlaIyAlaIyIleuAaPryIleProAaPryAlaTyrlle 120
Db 202 -----AGCTTAC--- 207
Qy 121 ValGlnTyrglnGlyAaPValGlnSerIyValArgSerIleGlnIhIsvAlGlnSerVal 140
Db 208 -----GAGACAGGAGCAGGCGGTTAAGTATGAGAGCTCATGCGGCGCCAG 255
Qy 141 GlnProTyrlleuProIyIyTyrlleIleAspProGlnIleuPheThr-----IyGlyAla 158
Db 256 GTCAAGTAC-----TCTTACAAAGATATATCCCTGCTGCGGTTAAATTAAGCCAGG 309
Qy 159 SerThrIleuValIyAaIeAaIeAaIeAaIeAaIeAaIeAaIeAaIeAaIeAaIeAaIe 176
Db 310 GACCTTCTGCTGATCGGCGGAGATATGACAGCGGTTACTTCGTTACACAAAGGTTCCG 369
Qy 177 GlnLeuArgGlyIleGlnIyIleAlaGlnTyValAlaIeAaenAaPValhIyTyrlle 196
Db 370 GGCAATAAGTTCAATACAGAG----- 390
Qy 197 ThrAlaIyProGlnTyrlleValMetAaAspValAlaArgGlyIle-----ValIys 214
Db 391 -----GATTCAAAGTTCAAGTTGACAGCCCACTTCCTCCAGATAGGG 438
Qy 215 AlaAspValAlaGlnSerSerTyrgIyLeuTyrgIyGlnIyGlnIyAlaIeAaIeAaIeAa 234
Db 439 GCGATACCGCTGCAACTCTCTCGGCTACGACGAGAGCGGTGCTGCTTCATATGTC 498
Qy 235 AspThrGlyLeuAaPThrGlyArgAaAspSerSerMetIhIsvAlaIeAaenGlyIyIys 254
Db 499 GATACGGGTATATACGCGAAG-----CACCCGATCTGAAGGCGAAG 540
```

QY	255	11ETRLALALEUYY---ALALEUylagYThRbAAsnAlaAspAspThRAsnGlyHs	273
Db	541	GTCAATAGCGCTGTACAGACCCCTCAACGGCAGAGTGAACCCCTTACAGATGACCAAGGAGCAC	600
QY	274	GLYTHRHLEVALALAGlySerValLeuGlyAAsnGlyAlaThRAsnLys-----Gly	290
Db	601	GGAAACCCACGTTGCGGGGTATCGTTCGCCGGAACCGGACAGGTTAACTCCACATGACATAGGC	660
QY	291	MetAlaProGlnAlaAsnLeuValPheGlnSerIleMet-----AspSerSerLysGly	308
Db	661	GTGCGCCCCGGCGCAAGCTCGTCGGAGCTCAAGGTTCTCGGTGCGGACGGTTCCGGAAGC	720
QY	309	LeuGlyGlyLeuProSerAsnLeuGlnThRLeuPheSerGlnAlaPheSerAlaGlyAla	328
Db	721	GTCTCCACCATCATGCGCGGTGTATCATGGGTCTGTCAGAAACAAAGACAAAGTACGGCATG	780
QY	329	ArgIle-----HisThrAsnSerTrpGly	336
Db	781	AGGGCTACCAACTCTCCCTCGGCTCTCCAGAGCTCCGACGGAAACCGACTCCCTCAGT	840
QY	337	AlaAlaValaAsnGlyAlaTyThrThRAspSerArgAsnValaAspAspTyValArgLys	356
Db	841	CAGGCGCTGCACACACGCTTGGGAGCC-----	867
QY	357	AsnAspMetThrIleLeuPheAlaAlaGlyAAsnGlyArgProAsnGlyGlyThRILEser	376
Db	868	---GGTATAGTATAGTCTGCGCGCGCGGCAACAGCGGCGGAACACCAACCGGTGCGC	924
QY	377	AlaAspGlyThRAlaLysAsnAlaIleThRValGlyAlaThRGlAsnLeuArgProSer	396
Db	925	TCACCCGCGCGCGGACGAGGTCTAAACCGTGGTGA-----	963
QY	397	PheGlySerTyAlaAspAsnIleAsnHisValaAlaGlnPheSerSerArgGlyProThr	416
Db	964	-----GTGACAGCAAGCAACAATCGCACTTCTCAGACGGGAGCCGAC	1011
QY	417	LysAspGlyAlaGlyIleLysProAspValaMetAlaProGlyThRTrpIleLeuSerAlaArg	436
Db	1012	GCGGACGGAAAGCTTCMAACCGGAATCTGCGCCCCCGGTGACATCATAGCCCCCGGC	1071
QY	437	SerSerLeuAlaProAspSerPheThRAlaAsnHisAspSerLysTyAlaTyMet	456
Db	1072	GCCAGC-----GGAACCAAGCATGGGACCCCGATMAACGACTACTACCAAGCC	1122
QY	457	GlyGlyThRSerMetAlaThRProIleValaAlaGly---AsnValaAlaGlnLeuArgGlu	475
Db	1123	TCTGGAACCAACATGCGACCCCGCATCTTTCGGCGGTGGCGGCGCTCATCTCCAGGCC	1182
QY	476	HisPheValLysAsnAspGlyLileThRProLys-----ProSerLeuLeuLysAlaAla	493
Db	1183	CAC-----CGAGCTGAGACCCCGGACAAAGTAAAGCCGCC	1218
QY	494	LeuIleAlaGlyAla-----AlaAspValGlyLeuGlyTyx	505
Db	1219	CTCATTCAGACCGCGCATAGTGGCCCCCAAGAGATAGCGGACATGCGCTTACGGTCGG	1278
QY	506	ProAsnGlyAAsnGlnGlyTrpGlyArgValThRLeuAspLysSerLeu-----Asn	522
Db	1279	-----GGTATGGGTGAACCTTCACAAAGGCGCATCAATACAGACGAC	1317
QY	523	ValAlaTyValaAsnGlySerSerAlaLeuSerThRserGlnLysAlaThRTrpPhe	542
Db	1318	TACGCCAAGCTCACTTACCGCGCTCGTCCGCCACAAAGGAGCGCACCAACCTTTC	1377
QY	543	ThRAlaThRAlaGlyLysProLeuLysIleSerLeuValTrpSerAspAlaProLaser	562
Db	1378	GACGTCAAGCGCGCACCTTGTGACCGCCACCTCTACTGGGAC-----	1422
QY	563	ThRThRAlaSerValThRLeuValAsnAspLeuAsnLeuValIleThRAlaProAsnGly	582
Db	1423	-----ACGGGCTCGAGGACATCAACCTTACTTACAGACCCCAAGCGG	1467

```

QY 583 ThrAGTVrVaIGLYAsnAhpheSerAAlProPheAsnAhpAntRrhpAsgLYArGAsn 602
Db 1468 AACGAG-----GTTGACTACTCTCCACACCGCTACTAC 150
QY 603 AsnValGIuAsnValPheIleAsnSerProGInSerGIYThrTYrThrIleGIuValGIn 622
Db 1501 GGCTTCGAGAAAGGTGGCTACTACAAACCGACCGCGGAACCTCGACGGTCAGGTCTGC 156
QY 623 AlaTYrAsnValProValGIYProGInAsnPheserIeAlaIleValAsn 639
Db 1561 AGCTACAAAG-----GGCGCGCGAACTACAGGTGACGTCTGTCAAGC 1602

RESULT 12
AAK05926
ID AAK05926 standard; DNA; 1977 BP.
XX
AC AAK05926;
XX
DT 06-MAY-1999 (first entry)
XX
DE WO9856926 Seq ID 11.
XX
KW Hyperthermostable; protease; thermophilic; bacterium; subtilisin;
KM additive; drug; washing agent; foodstuff; chemical synthesis; ds.
XX
OS Thermococcus celer.
XX
PN WO9856926-A1.
XX
PD 17-DEC-1998.
XX
PF 04-JUN-1998; 98WO-JP002465.
XX
PR 10-JUN-1997; 97JP-00151969.
XX
PA (TAKI ) TAKARA SHUZO CO LTD.
XX
PI Takakura H, Morishita M, Shimomojo T, Anada K, Kato I;
DR WPI; 1999-080907/07.
XX
P-PSDB; AAM94840.
XX
PT Recombinant hyperthermostable protease from Pyrococcus furiosus - and
PT gene encoding it, for large scale production of the protease for
PT industrial use.
XX
PS Disclosure; Page 53-54; 82pp; Japanese.
XX
XX
CC The invention relates to a hyperthermostable protease derived from a
CC thermophilic bacterium (especially Pyrococcus furiosus). The protease has
CC working temperature 40-110 deg.C (optimum 80-95 deg.C), working pH 5-10
CC (optimum 6-8), and retains more than 50% of its activity after 8 hours at
CC 95 deg.C. The invention also provides gene sequences encoding a
CC polypeptide of formula SIG-Ala-Gly-Gly-Asn-PRO, where SIG is a signal
CC peptide from subtilisin, and PRO is the above protease. Host cells
CC (especially Bacillus strains) transformed with vectors comprising the
CC genes are used for the recombinant production of the protease. The
CC hyperthermostable protease which can be prepared in quantity suitable for
CC industrial use, can be used as an additive for drugs, washing agents and
CC foodstuffs and for chemical synthesis
XX
SQ Sequence 1977 BP; 453 A; 659 C; 554 G; 311 T; 0 U; 0 Other;

Alignment Scores:
Pred. NO.: 3.32e-27 Length: 1977
Score: 477.00 Matches: 168
Percent Similarity: 41.8% Conservative: 90
Best Local Similarity: 27.2% Mismatches: 203
Query Match: 14.6% Indels: 156
Gaps: 25

US-10-784-870-4 (1-639) x AAK05926 (1-1977)

```

US-10-784-870-4 (1-639) x AAX05926 (1-1977)


```
Db 910 AAGTATACATCGTTCACAGCAGTCGACAGCAAGTATTACAGTTGAGCC----- 963
Qy 392 AATLeuArgProSerPheGlySerTyrAlaAspAsnIleAsnHisValAlaGlnPheSer 411
Db 964 -----GTTGACAGATGATGATTTAAACAGCTTCA 996
Qy 412 SerArgGlyProThrIleAspArgIleValProAspValMetAlaProGlyThrTyr 431
Db 997 ACGAGAGGCCCACTGACAGCGGACGCTTAAGCTGAGGTGTGCTCCAGGAAACTGG 1056
Qy 432 IleLeuSerAlaArgSerSerLeuAlaProAspSerPheThrAlaAsnHisAspSer 451
Db 1057 ATATATGCTGCCAGAGCAAGT-----GAACTAGCATGGGTCAACCAATTATATGAC 1107
Qy 452 LysTyrAlaIleArgMetGlyIleThrSerMetAlaThrProIleValAlaGlnAsnValAla 471
Db 1108 TATTACACAGCAGCTCTGGGACATCATGCAACTCTCAGCGTATGCGTATTCAGCC 1167
Qy 472 GlnLeuArgGlnHisPheValIleAsnArgGlyIleThrProLys-----ProSerLeu 489
Db 1168 CTCTTCTCTCA-----GACACACCCGAGCTGAGCTCAGACAA 1206
Qy 490 LeuValAlaAlaLeuIleAlaGlyAla-----AlaAspVal 501
Db 1207 GTAATAACAGCCCTCATAGAACTGCTGATATGCTTAAGCCAGATGAAATAGCCGATATA 1266
Qy 502 GlyLeuGlyTyrProAsnGlnIleValGlyTyrGlyArgValThrLeuAspIleSerLeu 521
Db 1267 GCTTACGCTGCA-----GTTAGGCTTATGATCATGACAGGCTATA 1305
Qy 522 AsnValAlaIleTyrValAsnGlnIleSerAlaLeuSerThrSerIleValAla----- 538
Db 1306 AAC-----TACATATACATGACAAAGCTAGTGTTCAGTATGTTGCCAACAAAGGC 1359
Qy 539 -----ThrTyrThrPheThrAlaThrAlaGlyLysProLeuValIleSerLeuValTyr 556
Db 1360 AGCGAACTCACACAGTTGCTTATAGCGAGCTTCGTTGTAATGCGACATTTATCTGG 1419
Qy 557 SerAspAlaProAlaSerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuVal 576
Db 1420 GACATGCCCCAT-----AGCACCTTGATCTTTAC 1449
Qy 577 IleThrAlaProAsnGlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsn 596
Db 1450 CTCTACGATCCCAATGAAACAG-----GTTGACTAC 1482
Qy 597 AsnTyrAspGlyArgAsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThr 616
Db 1483 TCTTACACCGCCTACATATGATTCGAAAGGTTGGTTATTACAACTGATGAGACA 1542
Qy 617 TyrThrIleGlnValGlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAla 636
Db 1543 TGGACATTAAGTGTATAGCTACAGC-----GAAAGTGCANATATCAAGTATGAT 1593
Qy 637 IleValAsn 639
Db 1594 GTGTAAAT 1602

RESULT 15
ACL64496/c
ID ACL64496 standard; DNA; 9125 BP.
XX ACL64496;
XX 02-JUN-2005 (first entry)
XX M. xanthus DNA fragment, seq id 959.
XX Transgenic plant; DNA replication; gene regulation; gene expression; de.
XX Myxococcus xanthus.
XX
```

```
PN US6833447-B1.
XX
XX 21-DEC-2004.
XX
XX 10-JUL-2001; 2001US-00902540.
XX
XX 10-JUL-2000; 2000US-0217883P.
XX
XX (MONS ) MONSANTO TECHNOLOGY LLC.
XX
XX Goldman BS, Hinkle CJ, Slater SC, Wiegand RC;
XX WPI; 2005-028716/03.
XX
XX New substantially purified Myxococcus xanthus nucleic acid molecule
XX encoding a nitrite reductase, useful for determining gene expression,
XX identifying mutations in a gene of interest, and for constructing
XX mutations in a gene of interest.
XX
XX Example 1; SEQ ID NO 959; 25pp; English.
XX
XX The invention relates to a substantially purified nucleic acid molecule
XX encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
XX recombinant DNA construct for expression of a nitrite reductase gene in a
XX plant cell, and a plant cell comprising the recombinant DNA construct.
XX The nucleic acid is useful for determining gene expression, identifying
XX mutations in a gene of interest, and for constructing mutations in a gene
XX of interest. Sequences given in records for SEQ IDs 1-1849 represent a
XX set of 1849 contig and singleton sequences comprising coding sequences,
XX CC DNA replication elements, promoters and other regulatory elements from
XX the genome of the bacterium Myxococcus xanthus. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO
XX
SQ Sequence 9125 BP; 1331 A; 2947 C; 3329 G; 1518 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 9.73e-24 Length: 9125
Score: 443.00 Matches: 172
Percent Similarity: 42.9% Conservative: 108
Best Local Similarity: 26.3% Mismatches: 231
Query Match: 13.6% Indels: 142
DB: 14 Gaps: 29
US-10-784-870-4 (1-639) x ACL64496 (1-9125)
Qy 60 PheLeuValGlnSerGlnAsnValIleValLeu-----IleValGly-----IleLeu 74
Db 8329 TACTGTGTGCGGCTGAGACACCGCGCTTGCACCGCTGAAGGGGCGACAGCCCTGGTC 8270
Qy 75 IlyValLeuGlnThrValProAlaAsnAsnIleValHisIleValGlnPheAsnGlyPro 94
Db 8269 GCGGGCGTGTACACCTCGGAGCGGACGAAACACCTGCACTGTGTGATTCACCGCAG 8210
Qy 95 IleLeuGlnGlnThrIleValGlnIleValGlnIleValGlnIleValGlnIleValGln 114
Db 8209 CCCCTGCGGAGTACCGGACCGCATCAAGCGCGCGGCGGAGATGCTGCGCTTCTTC 8150
Qy 115 ProAspTyrAlaTyrIleValGlnIleValGlnIleValGlnIleValGlnIleValGln 134
Db 8149 ACGAGACACACCTTCTGCTGAGATGATGATCCAGACGACGACGACGACGACGACGACG 8090
Qy 135 GluIleValGlnSerValGlnProTyrIleProTyrIle----- 147
Db 8089 CCTATATGCTCGGTGTGGGCGGTACATCCGAGTACCGGTGTGGGAGGCGGTGTGGGCG 8030
Qy 148 -----LysIleAspProGlnIlePheThr----- 155
Db 8029 GAGTCTTCAGGGGCGCTGCTGAGAGCTGAGAGCGGACGCTACTCATCATGCTGGGCG 7970
Qy 156 ---LysGlyAla-----SerThrLeuValIleValAlaLeuAlaLeuAsp 168
Db 7969 GAGCGGCGGCGGACGACGACGAGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7910
```



```

QY 169 ThrlyGlnAAsnValGlnLeuArglyLeu-----181
DB 7909 GTGAGCTC-----ATCAAGCGGGCGGCTCGCGGTGGAGCCACCTCCACC 7862
QY 182 ---GluGlnLeuAlaGlnTyValAlaSerAsnValAlaValTyrLeuThr-----Ala 198
DB 7861 CAGTCCCAAGTGGAGCGGTGATTCGCTCCAAAGCGCGTGCAGTACGACCGCTGGAGCG 7802
QY 199 LysPArglyTyValValMetAsnValAlaAlaGlyLeuValValAlaAspValAla 218
DB 7801 GGGCGGGGT---GAGCTGCAAAACAAACGTCGCGATGTGGGCGGCGCACTTACCTC 7745
QY 219 GlnSerSerTyGlyLeuTyGlyGlnGlyGlnLeuAlaValAlaAspThrGlyLeu 238
DB 7744 GAAAGCCTAAGAGGCTTCAAGCGGACAGGCGCTCCGCGGAAATCTTGACACCGAGGTG 7685
QY 239 AspThrGlyArgAsnAsp-----SerSerMetHisGlu 249
DB 7684 CGCACACGACCAAGAGTGGCGCACCTCCCTCATCCACAGACGTCACACGAGGAC 7625
QY 250 AlaPheArgGlyLeuLeuThrAlaLeuTyValAlaLeuGlyArgThrAsnValAlaAsp 269
DB 7624 ACCTTCCACGGT---ACCGCTGCTACAGCATC-----7595
QY 270 ThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGlyValAsnGlyAlaThrAsn 289
DB 7594 -----AACTTCGCGCGGGCGGTGACGGGAACGCG-----GCG 7562
QY 290 GlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerGlyLeu 309
DB 7561 GGCCTGTGCTCCCTCCGACAAAGCATCTTACCCGTACAAAGAGTCTCC---CAGTTC 7505
QY 310 GlyGlyLeuProSerAsnLeu-----GlnThrLeuPheSerGlnAlaPheSerAla 326
DB 7504 GGGGGGCGCAAGTCTCGTACCAATCAACCGGAGCTACGACCGCGCGCGGTAC 7445
QY 327 GlyAlaArgIleHisThrAsnSerTrpGlyAlaAlaValAsnGlyAlaTyThrThrAsp 346
DB 7444 CGGCGGTGTTCAGACGTCACAGCGTGGCAGCGCGCAGGTGACGCTACACCAATC 7385
QY 347 SerArgAsnValAspAspTyValAlaGlyValAsnAspMetThrIleLeuPheAlaGly 366
DB 7384 TCGCGAGAGTGTGATGATCACTGTTCACTACCC-----ATCCTCAAGACGCGAG 7334
QY 367 AsnGlyArgProAsnGlyGlyThrIleSerAlaPro---GlyThrAlaValAsnAlaIle 385
DB 7333 TCAAGAGCAACACCGCAACCAAGACTCGCGGCAAGCGGTGGCAAGAACATCTGT 7274
QY 386 ThrValGlyAlaThrGlnAsnLeuArgProSerPheGlySerTy---AlaAspAsnIle 404
DB 7273 TCGTGGGT-----GGCATGTACCACTACGACCAACACC 7241
QY 405 AsnHisValAlaGln-----PheSerSerArgGlyProThrLysAspGlyArg 420
DB 7240 AACCGCGGAGTACCGGTGAAATGGCGGCGGAGCTGGGTCCGGCGCGGACGCGG 7181
QY 421 IleLysProAspValMetAlaProGlyThrTyIleLeuSerAlaArgSerSerLeuAla 440
DB 7180 CTCAGCGCGAGCTGTGATCTACAGACCTATCGGCTCGCGGTGAAACACC-----7127
QY 441 ProAspSerPheThrAlaAsnHisAspSerLysTyValAlaTyMetGlyGlyThrSer 460
DB 7126 -----AGCAACACGTCCTACCAACTTCGGCGGACCGAC 7091
QY 461 MetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlnHisPheValLysAsn 480
DB 7090 GCGGCAAGCGCGAGAGCGGGGCACTTGGGCTGTCTTCCAGATGTG-----CAC 7037
QY 481 ArgGlyIle-----ThrProLysPro 487
DB 7036 GAGGCGCTGTGGCGGCTTCGGTGTGGCGGAGTGTTCGCAAGCGCTCGAAGATG 6977

```

```

QY 488 SerLeuLeuValAlaAlaLeuIle-----AlaGlyAla 498
DB 6976 GCCACGCGCAAGCGCGTGAATCAATGAGCCATCGTCAACATCGCTCGCGGCGCGC 6917
QY 499 AlaAspValIleLeuGlyTyProAsnGlyAsnGlyTyTrpGlyArgValThrLeu---517
DB 6916 TCCAGCTGATCACTG-----ACGCGCGCGGCAAGGCTGGGTACCGCGGACGTGAAG 6863
QY 518 -----AspLysSerLeuAsnValAlaTyValAsnGlySerSerAlaLeuSerThr 534
DB 6862 CGCGTGTATGACCGCGCGCGGTGACAGCATCATCAACAGACGCAACCTGTCTCCG 6803
QY 535 SerGlnLysAlaThrTyThrPheThrAlaThrAlaGlyLys---ProLeuLysIleSer 553
DB 6802 CTCGCGCAAGATACGTCAACGTCACTGCGCTCCCGGGAAGACGAGCTCAACGTACC 6743
QY 554 LeuValTrpSerAspAlaProAlaSerThrAlaSerValThrLeuValAsnAspLeu 573
DB 6742 ATGATCTACACCAACCGCGCGGACCGGTGGCGCGGCGCAACCGCATCAACGACTG 6683
QY 574 AspLeuValIleThrAlaProAsnGlyThrArgTyValAlaAsnAsp---PheSerAla 592
DB 6682 TCGCTGCGCGTGAAGTCCCGCACCGCGGTGCTTACGTGGGCAACAGCGCTGACGCGG 6623
QY 593 ProPheAspAsnAsnTrpAspGlyArgAsnAsn-----ValGlyAsnValPhe 608
DB 6622 AGCAACCTTCACAGCGCGGCGGTGTGAAACAGGTGATACGCTGAGAAAGTCTTC 6563
QY 609 IleAsnSerProGlnSerGlyThrTyThrIleGlyVal 621
DB 6562 CTCGCAACCGCGCGCGGCGGACGTGACCGGTGAGGTG 6524

```

Search completed: April 8, 2006, 03:44:56
Job time : 1707 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 8, 2006, 03:09:21 ; Search time 6632 Seconds

(without alignments)
4507.983 Million cell updates/sec

Title: US-10-784-870-4

Perfect score: 3267
Sequence: 1 MKKKVFLSVLAAALSTV.....EVOQNVVPGPNFSLATV 639

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_plus.model -DEV=xlh
-O=/abse/ABSEWEB.spool/US10784870/runat_07042006.143417_27495/seq_query.fasta_1
-DB=EST -QFMT=fastlap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abse08
-USER=US10784870.OCGN_1_1_5315@runat_07042006.143417_27495 -NCPU=6 -ICPU=3
-NO NMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAFN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsa1:*
10: gb_gsa2:*
11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	282.5	8.6	689	8	CV960494 PXrpxc 89
2	258	7.9	1605	5	BQ622771 CC Contig
3	246.5	7.5	627	10	CW807544 pine 1964
4	245.5	7.5	4398	4	CR861261 Pongo pyg
5	244	7.5	4198	4	AK029048 Mus muscu
6	243	7.4	3159	11	DD050980 Homo sapi
7	243	7.4	3159	11	DD050981 Pan trogl

8	242.5	7.4	679	10	CW807764	CW807764 pine 8939
9	231	7.1	540	3	BJ395336	BJ395336
10	220.5	6.7	594	3	BJ393752	BJ393752
11	219.5	6.7	1595	4	CNS0A30A	BX827771 Arabidops
12	219	6.7	962	8	DR651450	DR651450 EST104156
13	217.5	6.7	532	11	TA319610P	AL492464 T. brucei
14	217.5	6.7	874	7	CN808705	CN808705 EST0051 M
15	214	6.6	555	10	CN803334	CN803334 pine 1830
16	214	6.6	729	5	BM449844	BM449844 BM449844
17	213.5	6.5	839	7	CN809011	CN809011 EST0357 M
18	213.5	6.5	863	7	CN809003	CN809003 EST0349 M
19	213.5	6.5	2337	10	CL974130	CL974130 OeIFCC042
20	212	6.5	632	10	CN817721	CN817721 pine 906
21	212	6.5	2619	10	CL962002	CL962002 OeIFCC007
22	209	6.4	841	7	CR939585	CR939585 CR939585
23	208.5	6.4	2141	3	BQ142519	BQ142519 Contig M
24	208	6.4	2208	5	CL966076	CL966076 OeIFCC013
25	207	6.3	720	5	BM449533	BM449533 BM449533
26	205.5	6.3	826	7	CN809152	CN809152 EST0498 M
27	202.5	6.2	2202	10	CL979759	CL979759 OeIFCC044
28	202	6.2	939	8	DR602738	DR602738 EST92866
29	201.5	6.2	1771	4	CNS0A32Y	BX827822 Arabidops
30	201	6.2	508	9	AO652212	AO652212 Sheared D
31	200	6.1	1550	8	DR010221	DR010221 Uf043 hau
32	199.5	6.1	601	3	BJ387574	BJ387574 BJ387574
33	199.5	6.1	744	8	DR628193	DR628193 EST101832
34	198.5	6.1	707	8	DR628018	DR628018 EST101814
35	198.5	6.1	885	8	DR624615	DR624615 EST101474
36	198	6.1	797	8	DR657363	DR657363 EST104748
37	197	6.0	822	8	DR657362	DR657362 EST104747
38	196	6.0	633	3	BJ369190	BJ369190 BJ369190
39	196	6.0	690	8	DR624989	DR624989 EST101511
40	195	6.0	771	6	CA320325	CA320325 UI-M-FW0
41	195	6.0	1002	11	CNS0606B	AL93417 T3 end of
42	194.5	6.0	2274	10	CL960806	CL960806 OeIFCC037
43	193.5	5.9	824	7	CNS45097	CNS45097 UI-M-HN0
44	193	5.9	2472	4	CNS0A4W3	BX842380 Arabidops
45	192.5	5.9	659	5	BM448603	BM448603 BM448603

ALIGNMENTS

RESULT 1

CV960494

LOCUS 689 bp mRNA linear EST 25-JAN-2005
PXrpxc 8939 mycellium, starved in water Phytophthora infestans cDNA,
DEFINITION
CV960494
ACCESSION
CV960494.1 GI:58150285
VERSION
CV960494.1
KEYWORDS
SOURCE
ORGANISM
Phytophthora infestans (potato late blight agent)
Phytophthora infestans
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
Phytophthora.

REFERENCE
1 (bases 1 to 689)
Randall,T., Dwyer,R.A., Huitema,E., Beyer,K., Cvitanich,C.,
Kelkar,H., Fong,A.M., Gates,K., Roberts,S., Yazkan,E., Gaffney,T.,
Law,M., Testa,A., Torto-Alalibo,A., Zhang,M., Zheng,L., Mueller,E.,
Windasse,J., Binder,A., Birch,P.R.J., Gisel,U., Govers,F., Gow,N.A.,
Mauch,P., van West,P., Maugh,M.E., Yu,J., Boller,T., Kamoun,S.,
Lam,S.T. and Udelsdon, H.S.
Large-scale gene discovery in the oomycete Phytophthora infestans
reveals likely components of phytopathogenicity shared with true
fungi

JOURNAL
MOL. Plant-Microbe Interact. 18 (3), 229-243 (2005)
PUBMED
15782637

COMMENT
Contact: Judeleon HS
Department of Plant Pathology
University of California
Weber Hall, Riverside, CA 92521, USA
Tel: 909 787 4199
Fax: 909 787 4294
Email: howard.judeleon@ucr.edu.

FEATURES

```

Location/Qualifiers
1. 689
/organi sm="Phytophthora infestans"
/mol _type="mRNA"
/strai n="88069"
/db_xref="taxon:4787"
/sex="A1"
/clone _lib="mycelium"
/note="vector: pSFO11"

```

ORIGIN

Alignment Scores:	
Pred. No.:	2, 19e-20
Score:	282.50
Percent Similarity:	51.3%
Best Local Similarity:	36.5%
Query Match:	8.6%
DB:	8
Gaps:	10
Matches:	683
Conservative:	84
Mismatches:	79
Indels:	33
Gaps:	10

US-10-784-870-4 (1-639) X CV960494 (1-689)

231 ValAlaValAlaLeuArgHisGlyLeuMetPheTrpGly-ArgAsnAsp-SerSerMetHis----- 248
Db 3 ATAGGGAAGAGCTGTACTACGCTCTGACAGTACCGGATCCGGAATTCGCCGGTCCAGCCACCGCT 62
Qy 249 ----GluAlaPheArgGlyLeuValIleThrAlaLeuTrpAlaLeuGlyValArgThrAsnAsn 267
Db 63 CCGCGGACCGCTGGCGGCGGCGCCAAAGTCCGATGATACAAAT---GGCTGGGCAAAACAAG 119
Qy 267 IaAsnAsp-----ThrAsnGlyHisGlyThrHisValAlaGlySerValLeuG 283
Db 120 CGGATGAAGCCGAATCTATCACTCTGTGGCCAGCGATACAGTGGCTGGCTGTACTGT 179
Qy 283 IYAAsnGlyAlaThrAsnLeuS-----GlyMetAlaProGlnAlaAsnLeuValP 299
Db 180 GCAGCTCGTTTAGCGGTAAACATGCCAAATCTTGGCATTTGGATGCCAGACAGCATACAT 239
Qy 299 heGlnSerIleMetAspSerSerGlyGlyLeuGlyGly----- 311
Db 240 TTATGGACATCGGAGCCAGACCGAAACCTGCACAGGTCAAGTCAGTCAATTTGCCCTGTAC 239
Qy 312 ----LeuProSerAsnLeuGlnTrpLeuPheSerGlnAlaPheSerAlaGlyValArgI 330
Db 300 TAGCCACACTCTGCAGACGCCACAGCACTTGTCTAGAGATCAAAATTGACGTGTGTCCAA 359
Qy 330 IeHisThrAsnSerTrpGlyAlaAlaValaAsnGlyAlaIleThrThrAspSerArgAsn 350
Db 360 TTTTTCATCTTCGTGGGGGCAACACT---GGGTGCGATTACAGCTCAAGCCAGACAT 416
Qy 350 AlaAspAspIleValArgIleAsn---AspMetThrIleLeuPheAlaAlaGlyAsn--- 367
Db 417 TGGACGCTTTTATTACGAAGAACCGAAGATCTGTGTGTGTGTGGCGGCAAGCAAGTG 476
Qy 368 --GluArgProAsnGlyGly-----ThrIleSerAlaProGlyThrAlaIleAsnAlaI 385
Db 477 GCGAGCTTCGACCACTGGTCAAGCCCACTTATTCATCCATCCGGATGCCAAAAGCTAA 536
Qy 385 IeHisValGlyAlaThrGluAsnLeuAspProSerPheGlySerIle-----A 401
Db 537 TTTTGGTTGGCGATCGCTGAACCTGTGACGCTTCTTTCACGCAATTTGGTGGCTGGATG 596
Qy 401 IaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrIleAspGlyValArgI 421
Db 597 TATCCAAATGAGGCAACAGTTGATGATGTTTCTCTGCAGGGCCCAACAAGACGGAGAT 656
Qy 421 IeIleAspProAspValMetAlaProGly 429
Db 657 TAAAGCCGATGTAGTTGGCCGGAT 682

RESULT 2

LOCUS	1605 bp	mRNA	linear	EST
BQ622771				01-JUL-2002
CC_Config67	Candidobolus cornatus	ARSEP	512	Candidobolus coronatus

ACCESSION	BQ622771
VERSION	BQ622771.1
	GI:21649940

KEYWORDS

CDNA, mRNA sequence
BQ622771
BQ622771.1
GI:21649
EST.

**MOVIES
SOURCE**

CDN,
BQ6,
BQ6,
EST

SOURCE
ORGANISM

CDNA, mRNA sequence.
BQ622771
BQ622771.1 GI:21649940
EST.

ORGANIS

cdna, mRNA sequence.
 B0622771
 B0622771.1 GI:21649940
 EST.

CDNA, mRNA sequence.
B0622771
B0622771.1
GI:21649940
EST.

5555

REFERENCE

AUTHORS

TITLLE

JOURNAL COMMENT

Candidobolus coronatus during optimized secretion of proteins
Unpublished (2002)

Contact: Premosier F. M.
Department of Entomology
University of Maryland
4112 Plant Sciences Building, College Park, MD 20742, USA
Tel: 301 405 16 13
Fax: 301 314 92 90
Email: fff34@umail.umd.edu.

FEATURES

```

source
1. .1605
   /organism="Conidiobolus coronatus"
   /mol_type="mRNA"
   /strain="ARSEP 512"
   /db_xref="taxon:34488"
   /clone_1b="Conidiobolus cornutus ARSEP 512"
   /note="Vector: U12ap; Conidiobolus coronatus was grown in
minimal medium supplemented with Manduca sexta cuticle and
petcone for 18 hours. A cDNA library was constructed in
the unidirectional lambda vector U12ap."

```

Alignment Scores:	
Pred. No.:	4,67e-17
Score:	258.00
Percent Similarity:	43.1%
Best Local Similarity:	28.1%
Query Match:	7.9%
DB:	5
	22
Length:	1605
Matches:	108
Conservative:	58
Mismatches:	111
Indels:	108
Gaps:	22

Db 856 GATACGTGTCTCAATGTCAAGCCAAATGAC-----TTGGTGTGCT 897
Qy 255 TtethAlaLeuTyAlaLeuGlyArgTThraAsnAla-----AsnApThr 270
Db 898 GCCACT-----TGGGCTACTAACACGCTGTGTAGCAACCTGATGCT 942
Qy 271 AaagLYHieGLYThrHieValAlaAglySerValLeuGlyAaagLYAlaThrAenLYsgLY 290
Db 943 CACGGTCACGGTACTACTGCTGTGCTGCTACTATTCCTGCT-----ACCACCTATGGT 993
Qy 291 MetAlaProGlnAlaAsnLeuValPheGlnSerIleMet-----AspSerSerGlyGly 308
Db 994 GTTGCCAAAGAGCTTAACATTTGTCGGTTAAGCTTTAGTGATGATGCTCCGATCA 1053
Qy 309 LeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAla 328
Db 1054 TACTCTGGAAATATCTCCGGTATTTGACTGGGTTGTACAC-----TCTGCTGCCAAG 1107
Qy 329 ArgIleHieThrAsnSerTrpGlyAlaAlaValAaagLYAlaTyTrhTrhAspSerArg 348
Db 1108 AAGATTATCTCTATGAGTTTAGAGGTGTAGAACAGATGCTCTTAACCT----- 1158
Qy 349 AsnValAspAspTyAlaArgLYAsnAspMetThrIleLeuPheAlaAlaAglyAaagLY 368
Db 1159 GCTGTTAACACGCTCTCAGCAAGGAGTGTCTACTGCTGCTC-----CTGCTGTAAAGAT 1215
Qy 369 ArgProAaagLYGlyThrIleSerAlaProGlyThrAlaAlaValAaagLYAlaGly 388
Db 1216 AACAGAGATGCTTGTGATATCTCT-----CCGGCTTGTGCTCCCTCCGCAATACGTTGGT 1272
Qy 389 AlaThrGlu-----AsnLeuArgProSerPheGlySerTyAlaAsnAlaAsnHie 406
Db 1273 GCCACTGAGATCATATATAAAGGCTCATTTCTTAACCTCGTAGCTGTGCT----- 1326
Qy 407 ValAlaGlnPheSerSerArgGlyProThryAspGlyArgIleLYsProAspValMet 426
Db 1327 -----GATATCTTA 1335
Qy 427 AlaProGlyThrTyrlleuSerAlaArgSerSerLeuAlaProAspSerPheTrp 446
Db 1336 GCTCTGGTGTCAACATTCCTCCACC-----TGG 1365
Qy 447 AlaAsnHieAspSerSerTyAlaTyrlMetGlyGlyThrSerMetAlaThrProIleVal 466
Db 1366 AAGGATCTTAACCTGCCAACCAACACCATCTGTGATCCTTAAGGCTTGCCCTCAATT 1425
Qy 467 AlaGlyAsnValAla 471
Db 1426 GCTGCTTAGCTGCT 1440

RESULT 3
LOCUS CW807544 627 bp DNA linear GSS 20-JUN-2005
DEFINITION pline 1964 Phytophthora infestans T30-4 phytophthora infestans
genomic, genomic survey sequence.
ACCESSION CW807544
VERSION CW807544.1 GI:57981961
KEYWORDS GSS.
SOURCE Phytophthora infestans (potato late blight agent)
ORGANISM Phytophthora infestans
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
Phytophthora.
1 (bases 1 to 627)
RandalI.T.A., Dwyer,R.A., Huitema,B., Beyer,K., Cvitanich,C.,
Kelkar,H., Fong,A.M., Gates,K., Roberts,S., Yatskan,B., Gaffney,T.,
Lew,M., Testa,A., Torto-Alalibo,T., Zhang,M., Zheng,L., Wheller,B.,
Windaas,J., Binder,A., Birch,P., Gisel,U., Govers,F., Gow,N.A.,
Mauch,S.T., van West,P., Waugh,M.B., Yu,J., Boller,T., Kamoun,S.,
Lam,S.T. and Judelson,H.S.
Large-scale gene discovery in the oomycete Phytophthora infestans
reveals likely components of phytopathogenicity shared with true
fungi
Mol. Plant-Microbe Interact. 18 (3), 229-243 (2005)

PUBMED 15782637
COMMENT Contact: Judelson HS
Department of Plant Pathology
University of California
PLPA, Riverside, CA 92521, USA
Tel: 9518274199
Fax: 9518274294
Email: howard.judelson@ucr.edu
Class: shotgun.
FEATURES
source Location/Qualifiers
1..627
/organism="Phytophthora infestans"
/mol_type="genomic DNA"
/isolate="T30-4"
/db_xref="taxon:4787"
/sex="A1"
/clone_lib="Phytophthora infestans T30-4"
/note="Vector: pUC18"

ORIGIN
Alignment Scores:
Pred. No.: 1,94e-16 Length: 627
Score: 246.50 Matches: 77
Percent Similarity: 46.8% Conservative: 24
Best Local Similarity: 35.6% Mismatches: 72
Query Match: 7.5% Indels: 43
DB: 10 Gaps: 8
US-10-784-870-4 (1-639) x CW807544 (1-627)

Qy 269 AspThrAaagLYHieGLYThrHieValAlaAglySerValLeuGlyAaagLYAlaThrAsn 288
Db 12 GATTCACACGCTCACGGTACATTCGTTACGGGAATTATGTGTGGCCAA-----TTGGC 65
Qy 289 LYsGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerGlyGly 308
Db 66 TTGGAGTGGCACCGGAGCTACTGTGAGGATGTAAAGATGTCTCAAGGTTC 125
Qy 309 LeuGlyGlyLeuProSerAsnLeuGlnThrLeuPhe----- 320
Db 126 TATGGG-----TGGATTGCTAGCATGTTCCAGTTCATGTTGTCCTACGACA 176
Qy 321 -----SerGlnAlaPheSerAlaGlyAlaGlyIleHieThrAsnSerTrpGly 336
Db 177 CCCAGTGGACATCAAAAGACTGCTCAAGACCGAATATGTTAAACAGTTGGCGGT 236
Qy 337 AlaAlaValAaagLYAlaTyTrhTrhAspSerArgAsnValAspAspTyAlaArgLYs 356
Db 237 -----GGCGGTCAAGGCTCTCAAGATGTTGATGGGTTATCAACGCTGGCGAGCT 287
Qy 357 AsnAspMetThrIleLeuPheAlaAlaAglyAaagLYArgProAaagLYGlyThrIleuSer 376
Db 288 GCAAGGATCATTCCTGTCTAGCTCAGGAGATCTGCTTAACTCGCGTACTATTTGA 347
Qy 377 AlaProGlyThrAlaLYsAaagLYAlaThrValGlyAlaThrGluAsnLeuArgProSer 396
Db 348 TCGCTGGAGATTCCGCTAGTGTGATCAACGTTAGTGCGACCGAT----- 392
Qy 397 PheGlySerTyAlaAspAsnIleAsnHie-----ValAlaGlnPheSerSerArgGlyPro 415
Db 393 -----ATCAACGACGAGCTGCTCTTTCAGCAAGGCGCCG 431
Qy 416 ThrLYsAspGlyArgIleLYsProAspValMetAlaProGlyThrTyrlleuSerAla 435
Db 432 AAGAGTCAGAGTCTCCGAAACCGATGTGGACGCCAGGCTGTGATG----- 485
Qy 436 ArgSerSerLeuAlaProAspSerSerPheTrpAlaAsnHieAspSerTyAlaTyrl 455
Db 486 -----TCATCTGCTGACGAGAC-----GACGCTCTTACGCTTC 521
Qy 456 MetGlyGlyThrSerMetAlaThrProIleValAlaAglyAsnValAla 471

QY 552 -----11eserleu-ValTTP-----SerAspAlaProAlaSerThrThrAl 565
Db 1883 CGACGTGACGTGATGTCCTACATGTCGCCCTACTGTCTCCACGCTACTACTATGAGG 1942
QY 565 aserValThrleuValAsnAspLeuAspLeuVal11ethrAlaProAlaGlyThrArgTy 585
Db 1943 AATGCCAACATCGTTAT-----GTCCACATCCTCAATGCGATGGCGGT 1987
QY 585 rValGlyAsnAspPheSerAlaProPheAspAsnAsnTTP-----AspG1 600
Db 1988 CACAGAAAGAAATTTGGATTAAGCT-----GAGTGGCGAACCTTATTTACACAGAA 2038
QY 600 yArGAsnAsnValGluAsnValPhe11eAsnSer-----ProGlnSerGlyThr 616
Db 2039 TGGAGACACATGAAAGTGCGCTTCTCTACTCTCTCAGTGTGTGGCGCTGTGAGTTA 2098
QY 616 rTyThr11eGluVal 621
Db 2099 CCTTGCCATCTCCATT 2114
RESULT 6
LOCUS DQ050980 3159 bp DNA linear GSS 02-JUN-2005
DEFINITION Homo sapiens MBTPS1 gene, VIRUAL, TRANSCRIPT, partial sequence,
ACCESSION DQ050980
VERSION DQ050980.1 GI:66904179
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 3159)
Nielsen,R., Bustamante,C., Clark,A.G., Gianowski,S., Sackton,T.B.,
Hubisz,M.J., Piepel-Alon,A., Tanenbaum,D.M., Civeello,D.,
White,T.J., Sinsky,J.J., Adams,M.D. and Cargill,M.
Direct Submision
Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source 1..3159
Location/Qualifiers
1..3159
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
<1..>3159
/gene="MBTPS1"
/locus_tag="HC19002"
ORIGIN
Alignment Scores:
Pred. No.: 6.6e-15 Length: 3159
Score: 243.00 Matches: 140
Percent Similarity: 38.6% Conservative: 89
Best Local Similarity: 23.6% Mismatches: 224
Query Match: 7.4% Indels: 140
Gaps: 28
US-10-784-870-4 (1-639) x DQ050980 (1-3159)
QY 87 HisIleValGlnPheAsnGlyProIleLeuGluGluThrLysGlnLysLeuGluThrThr 106

Db 166 TAAATGTGGCTTTCATGATGATCTTTACAGCCAAAGCAGAAATTCATTTATTCAGT 225
QY 107 GlyAlaLys-----11leu 111
Db 226 GCCCTGAAGACGATGAAATGACAAATTTGAGATTCCTGAAACATCATCCAGT 285
QY 112 AspTyThrProAspTyThrAlaTyThrIleValGluTyGluGlyAspValGlnSerLysVal 131
Db 286 GACTACCTTGAATGATTTTGAAG--GTGATTCAGATTAAGAAAGAAAGAAAGCGGGCTG 342
QY 132 ArgSer11eGluHisValGluSerValGluProTyLysProTyLysTyThrLysIleAspPro 151
Db 343 CTAACTTGAAGATCATCTCAACATCAACGCGTCAACGCCCAAGAAAGCTTTGCT 402
QY 152 GlnLeuPheThrLysGlyAlaSerThrLeuValLysAlaLeuAlaLeuAspThrLysGln 171
Db 403 TCCCTC-----AGTATGCTGAATCTGACCCCAAGTACCTTCATGAAACCCGGTGG 456
QY 172 AsnAsnLysGluValGlnLeuArgGly11eGluGlu11eAlaGlnTyThrValAlaSerAsn 191
Db 457 AGCCAGAAAGTGGCAATCATCAGTCCCTGGCAAGCCAGCTTCCCTGGGCTCTGGC 516
QY 192 AspValHisTyThrIleThrAlaLysProGluTyThrLysValMetAsnAspValAlaArgGly 211
Db 517 TTCTGGCATCTACGGGAAGGCACTTGACACAGCGCTGTGAGACCATCCGCCCGCAG 576
QY 212 11e-----ValLysAlaAspValAlaGlnSerSerTyGlyLeuTyGlyGlnGly 228
Db 577 GTTGCCCAACACATCGACGACAGATGTGCTGG--CAGATGGAGATACAGGTGCTAAT 633
QY 229 GlnIleValAlaValAlaAspThrGlyLeuAspThrGlyAlaArgAsnAspSerMetHis 248
Db 634 GTAAAGTGTGCTTTTACACTGGGCTG-----AGCGAAGACAT 675
QY 249 GluAlaPheArgGlyLysThrIleThrAlaLeuTyThrAlaLeuGlyLysArgThrAsnAlaAsn 268
Db 676 CCCCACTTAATAAT-----GTAGAGAGAGAACCAATCGACCAAC 717
QY 269 -----AspThrAsnGlyHisGlyTyThrHisValAlaLysSerValLeuGly 283
Db 718 GAGCGAAGCTGAGACATGAGTGGGCAATGACATTCGTGGCAGGT--GTGATAGCC 774
QY 284 AsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeu--ValPheGlnSerIle 302
Db 775 AGCATGAGGAGATGCCAAGATTTGTCAGATGACGAATTCACATTTTCAGGGTCTT 834
QY 303 MetAspSerSerGlyLysLeuGlyLysLeuProSerAsnLeuGlnThrLeuPheSerGln 322
Db 835 ACCAATATTCAG-----GTATCTTACATCTGTGTTTGGAC 873
QY 323 AlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyGlyAlaAlaValAsnGlyAla 342
Db 874 GCCTTCAACTATGCACTTTTAAAGAAATGCAAGTGTAACTCAACATCGCGCGCG 933
QY 343 TyThrThrAspSerArgAsnValAspAspTyThr-----ValArgLysAsnAspMetThr 360
Db 934 GACTTCATGATCATCCGTTGTTCGACAGGTGGGAATTTAACAGCTAACATGATTAAC 993
QY 361 IleLeuPheAlaAlaGlyAsnGlyLysArgProAsnGlyTyThrIleSerAlaProGlyThr 380
Db 994 ATGGTTTCTGCTAATTCATGACAGCAAGCTCTTTATGCGACTGTGAATTAACCTGTGAT 1053
QY 381 AlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyThr 400
Db 1054 CAATGAATGATGATGAGATGAGCGGCATTTAC-----TTT 1089
QY 401 AlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGly-----TTT 414
Db 1090 GAAGATTAACATC-----GCCGCTTTTCTTCAAGGGGAATGACTACCTGGAGACTA 1140
QY 415 ProThrLysAspGlyArgIleLysProAspValMetAlaProGlyTyThrTyThrIleLeuSer 434

Db 1141 CCAAGAGCTACGATGTCATGAACCTGACATTGTC-----ACCTATGATGCTGAC 1191
Qy 435 A1AargSerSerLeuAlaProAbpSerSerPheThrAlaAsnHisAbpSerIleTyrrAla 454
Db 1192 GTCGGGGTTCTGGCGCTGAAGGGGGTGC-----CGG 1224
Qy 455 TymeG1yG1YThSerMetAlaThrProIleValAlaG1yAbnValAlaG1nLeuArg 474
Db 1225 GCCCTCTCAGGAGCAAGCTGTGCTTCTCCAGTGTGGAGTGGCTCTCACTTGTATAG 1284
Qy 475 GluH1sPheValIlePheAbnArgIleThrProIleProSerSerLeuLeuAlaAlaLeu 494
Db 1285 AGCACAGTCCAGAGAGGTGAGCTGTGTG-----ATCCCGCAGATGAAAGCAGGCCCTG 1338
Qy 495 IleAlaG1yAlaAlaAspValAlaG1yLeuG1yTyrrProAbnG1yAbn-----G1nG1y 511
Db 1339 ATCGGCTCAGCCCGAGGCTC-----CCCGGGGTCAACATGTTTGAGCAAGGC 1386
Qy 512 TrpG1yArgValIleLeu-----Aplys 519
Db 1387 CACGGCAAGCTGATCTGCTTGAAGCTTACAGATCTCAAGACTACAGCCACAGCA 1446
Qy 520 SerLeuAbnValAlaTyrrValAbnG1ySerSerAla-----LeuSerThrSer 535
Db 1447 AGTTTGAAGCCCACTACATAGATCTGAGTGTCCCTACATGTGGCCCTACTGCTCC 1506
Qy 536 GluIlyAlaThrTyrrPheThrAlaThrAlaG1yIlyProLeuIleIleSerLeuVal 555
Db 1507 CAGCCCATCTACTA-----GGAGAAATGCCG----- 1533
Qy 556 TrpSerAspAlaProAlaSerThrThrAlaSerValIleLeuValAbnAbpLeu 575
Db 1534 -----ACAGTGTATATGTCACATCTCACTCAAGCGCATGGAGTGC 1572
Qy 576 -----ValIleThrAlaProAbnG1yThrArgTyrrValG1yAbnAbpPheSerAla 592
Db 1573 ACAGAGAAATTGTGATTAAGCTGACTGCGCAGCCCTATTG----- 1614
Qy 593 ProPheAbnAsnAsnTrpAspG1yArgAbnAsnValG1uAbnValPheIleAsnSer--- 611
Db 1615 -----CCACAGAACGAGAC-----AACATGTAAAGTTGCCCTCTCTACTCTCG 1659
Qy 612 -----ProG1nSerG1yThrTyrrThrIleG1uVal 621
Db 1660 GCTTATGAGCCTGTGTCGGGCTACCTGGCCATCTCCATT 1698
RESULT 7
LOCUS DQ050981 3159 bp DNA linear GSS 02-JUN-2005
DEFINITION Pan troglodytes MBTSP1 gene, VIRUAL TRANSCRIPT, partial sequence,
ACCESSION DQ050981
VERSION DQ050981.1 GI:66904180
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Pan.
REFERENCE
AUTHORS 1 (bases 1 to 3159)
Meisen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeille,D.,
Hultes,T.J., Snihsy,J.J., Adams,M.D. and Cargill,M.
A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
(et) Plos Biol. 3 (6), E170 (2005)
JOURNAL
PUBMED 15869325
REFERENCE 2 (bases 1 to 3159)
AUTHORS Meisen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeille,D.,
Hultes,T.J., Snihsy,J.J., Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source location/Qualifiers
1..3159
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene
<1..>3159
/gene="MBTSP1"
/locus_tag="HC19002"
ORIGIN
Alignment Scores:
Pred. No.: 6.6e-15 Length: 3159
Score: 243.00 Matches: 140
Percent Similarity: 38.64 Conservative: 89
Best Local Similarity: 23.64 Mismatches: 224
Query Match: 7.44 Indels: 140
DB: 11 Gaps: 28
US-10-784-870-4 (1-639) x DQ050981 (1-3159)
Qy 87 HisIleValG1nPheAbnG1yProIleLeuG1nIleuThryG1nIlyLeuG1nThr 106
Db 166 TATATTGGCTTTCATATGATCTTATACAGCCAAAGCTAGAAATTCATTATTTCAGT 225
Qy 107 GluAlaIly-----IleLeu 111
Db 226 GCCCTGAAGACAGTGAAGTGAACAATTGAGAAATTATCTCGAAACAATCATCCAGT 285
Qy 112 AspTyrrIleProAbpTyrrAlaTyrrIleValG1uTyrrG1uIlyAbpValG1nSerIleVal 131
Db 286 GACTAACCTTACTGATTTTGTAG--GTGATTCAGATTAAGAAACAAACGAAAGCGGGCTG 342
Qy 132 ArgSerIleGluH1sValG1ySerValG1uProTyrrLeuProIlyTyrrIleAspPro 151
Db 343 CTAACTTGAAGATCATCCAAACATCAACAGGGCTCACGCCCCCAAGAAAGTCTTTCGT 402
Qy 152 GluLeuPheThrIlyG1yAlaSerThrLeuValIlyValAlaLeuAlaLeuAspThrIlyG1n 171
Db 403 TCCCTC-----AGTATGCTGAATCTGACCCCAAGACCTCGCATGAAGAACCCGGTGC 456
Qy 172 AsnAbnIlyG1uValG1nLeuArgIlyIleG1uIleAlaG1nTyrrValAlaSerAbn 191
Db 457 AGCCAGAAATGGCAATCATCATCCCTCGAAGACCCCTCTCGGGCTGCTGC 516
Qy 192 AspValHisTyrrIleThrAlaIlyProG1uTyrrIlyValMetAsnAspValAlaArgG1y 211
Db 517 TTCGGCATGCTACGGAGGCAATTCAGACAGACGCTGCTGAGACCATCCGCGCAG 576
Qy 212 Ile-----ValIlyAlaAspValAlaG1nSerSerTyrrG1yLeuTyrrG1nG1y 228
Db 577 GTTCCCAAGACATCGCAGGCGAGATGTCTGCG--CAAGTGGATATPACAGGTGAAT 633
Qy 229 GluIleValAlaValAlaAspThrG1yLeuAspThrIlyArgAbnAspSerSerMetHis 248
Db 634 GTAAGATGTGCTGTTTGTACACATGGCTG-----AGGAGAACATCAT 675
Qy 249 GluAlaPheArgG1yIlyIleThrAlaLeuTyrrAlaLeuTyrrArgThrAsnAsnAlaAsn 268
Db 676 CCCCACTTCAAAT-----GTGAAGAGACAACTGACCAAC 717
Qy 269 -----AspThrAsnG1yAlaG1yThrnIleValAlaG1ySerValIleG1y 283
Db 718 GAGCGAAGCTGATGATGGGTGGCCATGCGACATTCGTGGCAGGT--GTATAGCC 774
Qy 284 AsnG1yAlaThrAbnIlyG1yMetAlaProG1nAlaAbnLeu--ValPheG1nSerIle 302
Db 775 AGCATGAGGAGTGCAGAGATTTGCTCCAGATCAACAACTTCATTTTCAGGGTCTTT 834
Qy 303 MetAbpSerSerG1yIlyLeuG1yIlyLeuProSerAbnLeuG1nThrLeuPheSerG1n 322

```

Db      835 ACCAATATATGAG-----GTAATCTACACATCTGTTTGGAC 873
Qy      323 AAlaphSerAlaGlyAlaArgIleHsrThrsenSerTrpGlyAlaAlaValaAnglyAla 342
Db      874 GCCTTCACATATCCATTTTAAAGAGATTAACGTGTAAACCTCAGCATCGCGCCG 933
Qy      343 TTTThrThrsenSerArgenValaAspAspTyr-----ValArgSerAsnAspMetThr 360
Db      934 GACTTCATGATCATCCCTTGTGTGACAGAGTGTGGAAATTAACAGTAAACATATGATC 993
Qy      361 IleLeuPheAlaIleAlaAnglyArgProAnglyGlyThrIleSerAlaProGlyThr 380
Db      994 ATGGTTTCCTATGTCAGATGACGAGACCTTTATGAGCTCTGATATACCTGCTGAT 1053
Qy      381 AlAlaAsnAlaIleThrValGlyAlaIleThrGluAsnLeuArgProSerPheGlySerTyr 400
Db      1054 CAATGATGTGATTTGAGTGAAGCGGCGCATTTGAC-----TTT 1089
Qy      401 AlAlaAsnIleAsnHisValAlaIlePheSerSerArgGly----- 414
Db      1090 GAAGATTAATCATC-----GCCCGCTTTCTTCAAGGGAAATGACTACTGGAGACTA 1140
Qy      415 ProThrLysAspGlyArgIleLysProAspValMetAlaProGlyThrTyrIleLeuSer 434
Db      1141 CCAGAGGCTACGCTGCGCATGAACCTGACATTTCTC-----ACCTATGGTCTCGGC 1191
Qy      435 AlaArgSerSerLeuAlaProAspSerSerPheTrpAlaAsnHisAspSerLysTyrAla 454
Db      1192 GTCCGGCGGTTCTGCGTGAAGCGGGGTGC-----CGG 1224
Qy      455 TyrMetGlyGlyThrSerMetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArg 474
Db      1225 GCCCTTCAGGAGCCAGGTGTGCTCTCCAGTGTGGAGTGTCTGCTCACTTAATTAAGTG 1284
Qy      475 GluHisPheValIleAsnArgGlyIleThrProLysProSerLeuLeuValAlaLeu 494
Db      1285 AGCAGAGTCCAGAGAGGTGAGCTGGTG-----AATCTGCCAGCTATGAAGAGGCCCTG 1338
Qy      495 IleAlaGlyAlaAlaAspValGlyLeuGlyTyrProAsnGlyAsn-----Gingly 511
Db      1339 ATTCGCTCAGCCGCGAGGCTC-----CCCGGGGTCAACATGTTTGAGCAAGGC 1386
Qy      512 TrpGlyArgValThrLeu-----Aaplys 519
Db      1387 CACGGCAAGCTCATGCTCTCAGAGCCTATCAGATCTCAACAGCTACAGCCAGCGCA 1446
Qy      520 SerLeuAsnValAlaTyrValaAnglySerSerAla-----LeuSerThrSer 535
Db      1447 AGTTTAAGCCCCGACTACATAGTCTGAGTGAAGTCCCTACATGAGCCCTTACTGCTCC 1506
Qy      536 GlnLysAlaIleThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuVal 555
Db      1507 CAGCCCATCTACTAT-----GAGAGAAATGCG----- 1533
Qy      556 TrpSerAspAlaProAlaSerThrThrAlaSerValThrLeuValaAsnAspLeu 575
Db      1534 -----ACAGTGTGTAATGTCACTCACTCAACGCGCATGAGGAGTGC 1572
Qy      576 -----ValIleThrAlaProAnglyThrArgTyrValGlyLysAsnAspPheSerAla 592
Db      1573 ACAGAGAAATTTAATAAGCTTGAAGCTGAGCCCTATTTG----- 1614
Qy      593 ProPheAspAsnAsnTrpAspGlyArgAsnAsnValGluAsnValPheIleAsnSer--- 611
Db      1615 -----CCACAGAACGAGAGC-----AACATTGAAGTTGCTTCTCTACTCTCTCG 1659
Qy      612 -----ProGlnSerGlyThrTyrThrIleGluVal 621
Db      1660 GTCTTATGACCTTGTGCGGCTACCTGCGCATCTCCATT 1698

```

RESULT 8
CM807764

```

LOCUS      CM807764                679 bp    DNA    linear    GSS 20-JAN-2005
DEFINITION p1ne_89392 Phytophthora infestans T30-4 Phytophthora infestans
ACCESSION  CM807764
VERSION    CM807764.1
KEYWORDS   GI:57982181
SOURCE     Phytophthora infestans (potato late blight agent)
ORGANISM   Eukaryota; Eucaryotes; Oomycetes; Pythiales; Pythiaceae;
            Phytophthora.
REFERENCE  1 (bases 1 to 679)
AUTHORS   Randall,T.A., Dwyer,R.A., Huitema,E., Beyer,K., Cvitanich,C.,
            Kellar,H., Fong,A.M., Gates,K., Roberts,S., Yatzkan,B., Gaffney,T.,
            Law,M., Teeter,A., Torto-Alalibo,T., Zhang,M., Zheng,L., Mueller,B.,
            Windas,J., Binder,A., Birch,P., Gisl,U., Govers,F., Gow,N.A.,
            Mauch,F., van West,P., Waugh,M.B., Yu,J., Boller,T., Kamoun,S.,
            Lam,S.T. and Judelson,H.S.
            Large-scale gene discovery in the oomycete Phytophthora infestans
            reveals likely components of phytopathogenicity shared with true
            fungi. Plant-Microbe Interact. 18 (3), 229-243 (2005)
JOURNAL    Mol. Plant-Microbe Interact. 18 (3), 229-243 (2005)
PUBMED     15782637
COMMENT    Contact: Judelson HS
            Department of Plant Pathology
            University of California
            PUEB, Riverside, CA 92521, USA
            Tel: 9518274199
            Fax: 9518274294
            Email: howard.judelson@ucr.edu
            Class: hotgun.
FEATURES   Location/Qualifiers
            source          1..679
                                /organism="Phytophthora infestans"
                                /mol_type="genomic DNA"
                                /isolate="T30-4"
                                /db_xref="taxon:4787"
                                /sex="A1"
                                /tissue_type="total mycelium"
                                /clone_lib="Phytophthora infestans T30-4"
                                /note="vector: pUC18"
ORIGIN
Alignment Scores:
Pred. No.:      6,16e-16      Length:      679
Score:          242.50      Matches:      76
Percent Similarity: 46.7%      Conservative: 23
Best Local Similarity: 35.8%      Mismatches:  70
Query Match:    7.4%      Indels:      43
DB:             10      Gaps:      8
US-10-784-870-4 (1-639) x CM807764 (1-679)
Qy      273 HieGlyThrHisValAlaGlySerValLeuGlyAsnGlyAlaThrAsnLysGlyMetAla 292
Db      3 CACGGTACACACTTAACGGGAATTAATGCGGCCAA-----TTGCGCTTCGAGAGCGCA 56
Qy      293 ProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerGlyGlyLeuGlyLeu 312
Db      57 CCGGAGCTACCTGATGAGCATGTAAAGGATGTCGTCGCTCAAGGTTGCTATGCG----- 110
Qy      313 ProSerAsnLeuGlnThrLeuPhe----- 320
Db      111 ---TCGATTTGCTAGCATGTTTCAAGTTCAATGTTGTGCTTACGACACCCAGTGGACA 167
Qy      321 SerGlnAlaPheSerAlaGlyAlaArgIleHsrThrsenSerTrpGlyAlaAlaValaAsn 340
Db      168 TCAAAAGACTGCTCCAAAGCAGCAATTTGTTAAACACAGTTGGGT-----GGC 218
Qy      341 GlyAlaTyrThrThrsenSerArgenValaAspAspTyrValArgLysAsnAspMetThr 360
Db      219 GGTCAAGGCTCAGCATGTTTATGAGGGTTATCAACGCTTGGAGCTGACGAGATCATTT 278
Qy      361 IleLeuPheAlaIleAlaGlyAsnGlyArgProAnglyGlyThrIleSerAlaProGlyThr 380

```

Df	279	CCTGTCGTACGTCGCAGGGAAATCTGGTCTCTTAACCTGGGATCAATTGCGCATCGCTCGAAGAT	338
Oy	381	AAtayAsanAlalIethrValglYAlATnrgluAmleuArgProSerPheGlYserTy	400
Df	339	TCCGCTAGTGATATACGCTAGCTAGCTGACCAT-----	371
Oy	401	AlAapraenIlAasnHs---ValAlagInPheSerSeraGlYProThrLyasaglY	419
Df	372	-----ATCACACACGCCAGTCTCCTTTACACAGCAAAGCCCGCACAGTGCAGAGT	422
Oy	420	ArglieLysPProApvAlMeclAProglYlThrYrlleLeuSerAlarGerSerleu	439
Df	423	CTCCGAHAACCGGAATGGCGACCCCGAGGTGCTCTGGATATG-----	464
Oy	440	AlAProApSerSerPheTrylalaSnHAsaSerLyryALeTyMeGlYglYthr	459
Df	465	-----TACTTTCGTCGACGAC--GACGCTCTTACCTGCTTCAATCCGGCTCC	512
Oy	460	SerMetAlarThrProileValAlaglyAsnValAla	471
Df	513	AGTATGGCTGCACCCGACGCTCGCAGAGTGCATCGCT	548
RESULT 9			
Bj395336			
LOCUS	Bj395336	640 bp	mRNA linear EST 08-MAR-2002
DEFINITION	Bj395336 Dictyostelium discoideum cDNA library, SF Dictyostelium		
ACCESSION	Bj395336		
VERSION	Bj395336.1 GI:19306422		
KEYWORDS	EST.		
SOURCE	Dictyostelium discoideum		
ORGANISM	Dictyostelium discoideum		
REFERENCE	Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.		
AUTHORS	Unshihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.		
TITLE	Pull length cDNA of Dictyostelium discoideum at the slug stage		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: Tadasu Shin-I Center For Genetic Resource Information National Institute of Genetics 111 Yata, Mishima, Shizuoka 411-8540, Japan Tel.: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshin@genes.nig.ac.jp. Location/Qualifiers		
FEATURES			
Source	1..640		
	/organism="Dictyostelium discoideum"		
	/mol_type="mRNA"		
	/strain="AX4"		
	/db_xref="taxon:44689"		
	/clone="d638b16"		
	/sex="mat A"		
	/dev_stage="Slug stage"		
	/clone_lib="Dictyostelium discoideum cDNA library, SF"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.07e-14	Length:	640
Score:	231.00	Matches:	68
Percent Similarity:	50.0%	Conservative:	38
Best Local Similarity:	32.1%	Mismatches:	70
Query Match:	7.1%	Indels:	36
DB:	3	Gaps:	10
US-10-784-870-4 (1-639) x Bj395336 (1-640)			
Oy	224	LeutrygLyngLyngIlleValAlaValAlAapThrGLyleuApThrGLyArg--	242
Df	22	TTAAAGAGTTAAAGGCACATATGTGAGTATGCTGATACGTGTTTAGATGATGACCATTTG	81
Oy	243	-AsnaPse-----SerMethAgLuAlaPhaArgGLy	253

Db 82 TTCTTTGAGATTCAAAAGTATCCATATACATTATTAACAGTAATGAAAAATCATAGAAAA 141
 Qy 254 TysIIethrIaleuTyralaleuGIa-gThrAasnAlaasnaphThraengIYHIs 273
 Db 142 GTTGTAACT-----TATATTACTTACATGACAAATGAAGATTATGTAAATGCTCAT 192
 Qy 274 GIYTHIHIIeValAlaGIySerValleuGIyaengIY-----Ala 286
 Db 193 GGTACACAGATTGTGGCTCTCAGACAGGATCTCCAGAGATTCCTCATGGCTATTTCa 252
 Qy 287 ThrAenIyGlyMeclAlaProIInIlaenIeValPheGIInserIleMeclApsSer 306
 Db 253 TCAATTAATGTCCTTGCAACTGACGCAAGAAATGCATTTATGATCTT-----TCATCT 306
 Qy 307 GIYGIyLeuGIyGIyLeu---ProSerAenIeugInThrIeupheserGIInAlaPheSer 325
 Db 307 GGAAGTTCGAACCAACACACCACCAAGTTTCACTCAAAATGACAAACCATTAATATGAT 366
 Qy 326 AlAGIyAlaArgIleHieThraSerTrpGIyAla-----AlaValaengIY 341
 Db 367 GCAGGTGCAGAAAGTCATGTCGATTCCTGGGGTCTCGATCTTGCAGAGTATTAATGTGT 426
 Qy 342 AlATyTrhThrApsSerArgAsnValaApsPArTyValaArgIys---AsnApsMeclThr 360
 Db 427 GGTATATCCGATGATGAGTCGTGTCGTATGTATGATTCCTTATAGTACCAACCAATTCCT 486
 Qy 361 ILeleupheAlaAlaGIyAsnGIuArgProAengIyGIYThrIleSerAlaProGIYThr 380
 Db 487 ATACTTAAGAGCTGCTGTATTAACGAG---CAATTTGCATCTTTATTAATGCTCAAGCAACA 543
 Qy 381 AlAlYAsnAlaIleThrValGIyAlaThrGIuAenIeValArgProSerPheGIySer--- 399
 Db 544 GCTAAAAATGCATATTAACAGTTGTGCTCTGCAAAACAGCTCATGTAAATTAATGTGCAGAT 603
 Qy 400 -----TyralaApsAenIleAen 405
 Db 604 GCATTGGAATTAATTAATGATTTCTCAGATTAATGCTAAT 639
 RESULT 10
 BJJ393752 594 bp mRNA linear EST 08-MAR-2000
 LOCUS BJJ393752 Dictyostelium discoideum cDNA library, SP Dictyostelium
 DEFINITION dictoideum cDNA clone dds32b16 5', mRNA sequence.
 ACCESSION BJJ393752
 VERSION BJJ393752.1 GI:19304838
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum
 ORGANISM Dictyostelium discoideum
 BUKARYOTA; MYCETozoa; Dictyostelid; Dictyostelium.
 REFERENCE 1 (bases 1 to 594)
 AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.
 TITLE Full length cDNA of Dictyostelium discoideum at the slug stage
 JOURNAL Unpublished (2002)
 COMMENT Contact: Tadasu Shin-i
 Center for Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.
 FEATURES
 source
 1..594
 Location/Qualifiers
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="dds32b16"
 /sex="mat A"
 /dev_stage="Slug stage"
 /clone_lib="Dictyostelium discoideum cDNA library, SP"

Pred. No.:	141e-13	Length:	594
Score:	220.50	Matches:	54
Percent Similarity:	51.4%	Conservative:	40
Best Local Similarity:	29.5%	Mismatches:	64
Query Match:	6.7%	Indels:	25
DB:	3	Gaps:	7

US-10-784-870-4 (1-639) x BJ939352 (1-594)

QY	254	LYSILETHRALALEUTYRLALALEUGLYAGTRHRAENALAEANLAPHRASNGLYHS	273
DB	21	AAAGGTGAACCTATATTACCAACATCAACACGACGATGATTAAGTGAAGTGCAC	80
QY	274	GLYTHRHVALALAGLYSERVALLEUGLY-----ASNGLYALA	286
DB	81	GGTACACATATTATGGTCTTCAGACAGGTACTCCAGAGATTTCTTCAGTAAATTTTCA	140
QY	287	THRAENLYGLYMEALAPROGLNLAENLVALPHEGLNSERILEMECASPESER	306
DB	141	TCATTAGTGGCTTGGCTTCAGATCAAGATTCATTTGATTTGGCAAGTGGTTCA	200
QY	307	GLYGLYLEUGLYGLYLEUPROSERANLEUGINTHRLAEPHESERDIALAPHESEALA	326
DB	201	TCAATTTGACACCT---CCATCGAATTTGAAACATTTATCAACCATTTATGACGA	257
QY	327	GLYLAARGLILEHLETHRASERTRPGLYALA-----ALAVASNGLYALA	342
DB	258	GGTGCAAGATGCAATTGATGATTTCTGGGGTCTGATGATGAGAGGGGATACAGTGT	317
QY	343	TYTRHTRHAPRSEARHABENVALAPHAPRYVALALGLYVABN---APMETHRLLE	361
DB	318	TATTCATGACGACACGCTTCATTTATGATGATTCCTTTCCATCCATCCAGATTTTCATCAT	377
QY	362	LEUPHEALAEALAGLYASNGLYUARPROABNGLYGLYTHRLIASERLAPROGLYTHRALA	381
DB	378	CTTAGAGCTGCTGGTGAACAGACGAAATCTTATCACTACTACT---CAATCCACTGA	434
QY	382	LYBAENALAELETHRVALGLYALATHRGUBENLEUARG-----	394
DB	435	AAGAATGTATTAATTCGCTGGTGTCTCATCAAAACATTCATGAAATATTATTAACTGATGT	494
QY	395	PROSERPHEGLYSETRHIAAPHANLEAEN-----HIVALALAEINPHE	410
DB	495	CCAAATTTATTAATTAATTCATCATCTGCTATTAATTAACAGACTTAATATGATTTTC	554
QY	411	SERSEARX 413	
DB	555	GATAGCAGA 563	

RESULT 11
CNS0A30A
LOCUS
DEFINITION
Arbidiopsis thaliana Full-length cDNA Complete Sequence from clone GS1RPH1209 of Hormone Treated Callus of strain col-0 of
Arbidiopsis thaliana (thale cress).
BX827771
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arbidiopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurooids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1695)
Castaldi,V., Aury,J.M., Tallon,O., Winkler,P., Clepet,C., Menard,M., Craud,C., Queller,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
The Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL
REFERENCE
AUTHORS
Unpublished
2 (bases 1 to 1695)
Genoscope.

[illegible]

```

Db      466 GGGATGGGAAATGACATGTTTATGCCAAAAGCCAAATCGATCCGCTGTAGTGTGTT 525
Qy      379 G|YThrAlaIyAsnAla-----:::|
Db      526 GATGCTGCTAAGAACACAGACATAGTACTTGGCTAGTATGTTTCTCTGATCTTTG 585
Qy      385 -----IleThrValG|ValAlaThrGlu 391
Db      586 GATCGAAAAGAGTGAAGGAAAGGTATGTGTGTAGAAATGGAGAGAGTGTGTGGAG 645
Qy      392 AenIeuAargProSerPheGly-----SerTyrAlaAap 402
Db      646 TCTACTATCAAAAGCTATGAGAGTGTGTGTGTCATCATTTGATTAATATATCTTGAC 705
Qy      403 Aen-----:::|
Db      706 AATGCTCAATTTTCATGGCAGCTGCGACCAAGTTTAATAGCTCGGTGGCATATATATC 765
Qy      404 -----IleAen----- 405
Db      766 TACCGATATATCAACTCCACAAGATCAGCGTGGCTGTATTTCAGAAAATCGGCAAGTG 825
Qy      406 -----HisValAlaGlnPheSerSerArgGlyProThrIyAsnGlyArg 420
Db      826 ACAATCCCTGCTCCATTTGTGCTTCTTTTCATCAAGAGTCCC---AATCCGGATCA 882
Qy      421 Ile-----LysProAspValMetAlaProGlyThrTyrIleLeuSerAla----- 435
Db      883 ATACGCTTTTCAAGCTGATATCGCTGCACCGGAGTATGATATATGTGGCGCTTAATCT 942
Qy      436 -----ArgSerSerIeuAlaProAspSerSerPheThrAlaAsnHisAspSer 451
Db      943 CTAAAGAGATTAATGATGCTGGTTAGATGTGACACCCAGTTC-----TCA 987
Qy      452 LysTyrAlaIyAsnGlyIleThrSerMetAlaIleProIleValAlaGlyAsnValAla 471
Db      988 AATTTACCATCTCTGTTGGACCTCAATGGCTGCCCTCATGTTCTGTGTAGCTGG 1047
Qy      472 GlnIeuArgGlnIleIlePheValIyAsnArgGlyIleThrProIyAspSerIeuLeuIy 491
Db      1048 -----TACGTCAGAGTCTTTTTCATCCGAGATTTGACCCCTGCTGCACAA 1092
Qy      492 AlaAlaIleuIleAlaIyAlaIleAspValGlyLeuGlyTyrProAsnGlyAsnGly 511
Db      1093 TCGGCATATTAATCTCAGCAAAAACGATTAAGCCGG----- 1128
Qy      512 TrpGlyArgValThrIeuAspIySerIeuAsnValAlaTyr-----ValAsn 527
Db      1129 -----AGAGTGAACAAAGACGACAGATTTGCTTATGAGAGAGCCAAATTAAC 1176
Qy      528 GluSerSerAlaIeuSerThrSerGlnIyAlaIleThrTyr-----PheThr 543
Db      1177 CCACGACGAGCCGACAGCCCTGAGCTTAGTACACATGACGACATCTCTATGTCAGT 1236
Qy      544 AlaThrAlaGlyIyS-----ProIeuIyIleSerIeuValTrpSerAspAla 559
Db      1237 TCTTGTGGCGGAGGCTACACGACCACTAGCTCATTTGG---TGGACACGCT 1293
Qy      560 ProAlaSerThrThrAlaSerValThrIeuValIyAsnAspIeuAspIeu---ValIleThr 578
Db      1294 CCGTGAGCT-----GTTCCCTCCATTTGCTCCGTCGAGCTGGGCGACAG 1332
Qy      579 AlaProAsnGlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTrp 598
Db      1333 ATTCCTCACTACCTCCAAACATCACTCACTGAGATCCGCAAAACCTCCACATTTGG 1392
Qy      599 --AspGlyArgAsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrT 618
Db      1393 CTGTGTTCACGGCGAGAGTCAACACGTG-----GACACACCTGCTCGGTCTACA 1443
Qy      618 hrIleGluValGlnAla-----TyrAsnValProValGlyProGlnAsnAspSerI 655

```

```

Db      1444 CCGCACCGTCCGAGCACCGAAAGAGTGAATATCAGGTGAGACACAGATTTGTCAT 1503
Qy      635 euAla 636
Db      1504 TTCA 1508

RESULT 12
LOCUS   DR651450/c 962 bp mRNA linear EST 12-JUL-2005
DEFINITION EST1041567 FVN Gibberella moniliformis cDNA clone FVN015, mRNA
sequence.
ACCESSION DR651450
VERSION   DR651450.1 GI:70739926
KEYWORDS EST.
SOURCE   Gibberella moniliformis
ORGANISM Gibberella moniliformis
          Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
          Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE 1 (bases 1 to 962)
AUTHORS   Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y.,
          Utecher, R., Shultz, S., Feldblyum, T., Glenn, A.B., Plattner, R.D.,
          Kendra, D.F., Town, C.D. and White, J.A.
          Analysis of 87,000 expressed sequence tags reveals alternatively
          spliced introns in multiple genes of the fumonisin gene cluster
          unpublished (2005)
JOURNAL   Contact: Brown, D.W.
COMMENT   USDA/ARS/NCAR
          USDA
          1815 N. University St, Peoria, IL 61604, USA
          Tel: 309 681 6230
          Fax: 309 681 6689
          Email: brown@ncar.usda.gov
          TIGR sequence name: FVN015TV
          Seq primer: GTA ATA CGA CTC ACT ATA GGG C.
          Location/Qualifiers
            1..962
              /organism="Gibberella moniliformis"
              /mol_type="mRNA"
              /strain="m3125"
              /db_xref="taxon:117187"
              /clone="FVN015"
              /tissue_type="mycelia"
              /clone_id="FVN"
              /note="Vector: pBluescript II SK(+); XR; Site 1: EcoRI;
              Site 2: XhoI; anamorph: Fusarium verticillioides. Library
              FVN was obtained from RNA derived from a corn meal medium
              culture of strain M-3125. These cultures were prepared by
              inoculating an autoclaved mixture of 25 g corn meal and 5
              ml distilled water with 5 ml of water containing 5 x 10e7
              conidia. The inoculated medium was mixed thoroughly,
              distributed equally into two 100-mm plastic petri dishes,
              and incubated at room temperature for four or six days."

ORIGIN
Alignment Scores:
Pred. No.: 4,53e-13 Length: 962
Score: 219.00 Matches: 87
Percent Similarity: 44.4% Conservative: 40
Best Local Similarity: 30.4% Mismatches: 109
Query Match: 6.7% Indels: 50
DB: 8 Gaps: 15

US-10-784-870-4 (1-639) x DR651450 (1-962)
Qy      233 ValAlaAspThrGlyIleuAspThrGlyArgAsnAspSerSerMetHisGluAlaPheArg 252
Db      951 GTCTGACACACTGTGTTCAAC-----ACCAATATGAAATTTGAG 910
Qy      253 GlyIyIleThrAlaIeuTyrAlaIeuGlyArgThrAsnAsnAlaAsnAspThrAsnGly 272
Db      909 GGCCTGCTCAGCGCGCTTTATATCTGCTTCATGAGTGAACAAGCC---GATGTGTGGC 853
Qy      273 HisGlyThrHisValAlaIySerValIeuGlyAsn-----GlyAlaThrAsnIyGly 290

```



```

Db      852 CACGGTACTACGCTCTGGGACCATTCGGCGCAAGACCTACGGTGTCTCCAAAGGCC 793
Qy      291 MetAlaProGlnAlaLeuValPheGln---SerIleMetApsSerSerGlyLeu 309
Db      792 ACCATC---CAAGCTGTCAAGGTCTTCAGGCGCTCTTAGACACCTCATCATCTT 736
Qy      310 GlyIleuProSerAenLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAlaArg 329
Db      735 GCTGGTTTCAACGCGGTGCGCAAGATATCATCTCCAAAGGCCGCAACCTGCTCCGTC 676
Qy      330 IleHisThrAenSerTrpGlyAlaAlaValaAenGlyAlaTrpThrApsSerArgAen 349
Db      675 GTGAAC-----ATGCTCTGGGTGTGTGTTACTCTGCTTCTTCAACAC 631
Qy      350 ValApsAerGlyValAlaGlyAenApsMetThrIleuPheAlaAlaGlyAenGlyArg 369
Db      630 GCTGTCAACTCTGCTCTCATCTGATCATCTCTCCATTCCTGCTGGAAGCATGCT 571
Qy      370 ProAenGlyGlyThrIleSerAlaProGlyThrAlaAenAlaIleThrValAla 389
Db      570 GCCAAGCGCCGCAACACTCT---CTGGCTGTGTACCAAGCTATCACCGTGGTCT 514
Qy      390 ThrGlyAenLeuArgProSerPheGlySerTrpAlaApsAenIleAenIleValAlaGln 409
Db      513 ATGACAGC-----AGCTGGGC-----ATTGCTCT 487
Qy      410 PheSerAerGlyProThrIleApsGlyArgIleApsProApsValMetAlaProGly 429
Db      486 TACTCTTAACACGCGCACCGTCTT-----CACACTTGTGCTCCGGT 445
Qy      430 ThrTrpIleuSerAlaApsSerSerLeuAlaProApsSerPheTrpAlaAenHis 449
Db      444 ACCGCGGTCTCTGCTC-----TGTACACTAAC 415
Qy      450 ApsSerLeuTrpAlaTrpMetGlyGlyThrSerMetAlaThrProIleValAlaGlyAen 469
Db      414 AACAGCGCCCAACAGCATTAACCGTACTCCATGCGCCACTCCCATGCTGCTGCTT 355
Qy      470 ValAlaGln---LeuArgGlnHisPheValApsAenArgGlyIleThrProApsProSer 488
Db      354 GTCCTTACCGCATCTCTGTCAACGAGGTCTCC-----GGCGTCACTGGCGTACCAAC 301
Qy      489 LeuLeuValAlaAlaLeuIleAlaGlyAlaAlaApsValGlyLeu---GlyTrpProAen 507
Db      300 TGGCTCAAGACCATCTGCTACTCTGCGAAGATCATCGCAACCTCCGCACTCCCAAC 241
Qy      508 -----GlyAenGlnGly 511
Db      240 CTGATCGGCACCAACGCG 223

RESULT 13
LOCUS   TA119G10P          532 bp      DNA          linear      GSS 13-DEC-2000
DEFINITION
T. brucei sheared genomic DNA clone 119g10, forward sequence,
ACCESSION
AL492464
VERSION  AL492464.1  GI:11867408
KEYWORDS
GSS.
SOURCE   Trypanosoma brucei
          Trypanosoma brucei
          Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
          Trypanosoma.
REFERENCE
1 (bases 1 to 532)
  Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,
  Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
  Melville, S.B., Rajandream, M.A. and Barrell, B.G.
  Direct Submission
  Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
  project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
  Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
  nh@sanger.ac.uk
COMMENT
  Constructed at the Institute for Genomic Research (TIGR),

```

```

Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TRB927/4 GTRat 10.1) was mechanically sheared
to give a tight size distribution (
4 kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/projects/T_brucei/.

FEATURES
     source
         1..532
            /organism="Trypanosoma brucei"
            /mol_type="genomic DNA"
            /strain="TRB927"
            /db_xref="taxon:5691"
            /clone="319g10"

ORIGIN
Alignment Scores:
Pred. No.:      2,54e-13      Length:      532
Score:          217.50      Matches:      60
Percent Similarity: 48.4%      Conservative: 29
Best Local Similarity: 32.6%      Mismatches: 68
Query Match:    6.7%      Indels:      27
DB:             11      Gaps:          6

US-10-784-870-4 (1-639) x TA119G10P (1-532)

Qy      298 ValPheGlnSerIleMetApsSerSerGlyLeuGly----- 310
Db      2  GATATGCCAAGTAAATGATGTCGCCAGGCGGAAGATTCTTCAAGGGTGGCTGCC 61
Qy      311 -----GlyLeuProSerAenLeuGlnThrLeuPheSerGlnAlaPhe 324
Db      62 CATCCAGTCAAGACTTGTCTCCGCCACAGCTTACTCAATATTATTCGCCGGATAT 121
Qy      325 SerAlaGlyAlaArgIleHisThrAenSerTrpGlyAlaAlaValaAenGlyAlaTrp 344
Db      122 GGGCTGAGAGCCGCTGTGTTCTCAACTCGTGGGGTTTGTGCTCCCTCCAGTATTC 181
Qy      345 ThrApsSerAenAenValApsAerGlyValAlaGly---ApsApsMetThrIleLeuPhe 363
Db      182 GCTGTGAAAGATATGATGATGATGTCGCAAGTATATGACATGCGCTACTTATCTTC 241
Qy      364 AlaAlaGlyAenGlyArgProAenGlyThrIleSerAlaProGlyThrAlaApsAen 383
Db      242 TCCACTGGCAACGATTCACAGATGCG-----CTATGACTCTGCTGCTGTAAAGAC 295
Qy      384 AlaIleThrValGlyAlaThrGlnAenLeuArgProSerPheGlySerTrpAlaApsAen 403
Db      296 GTGATGTGGTGGCGGTCACACAAACATG-----TTTGACGCTTCGAAAGAC--- 343
Qy      404 IleAenHisValAlaGlnPheSerSerArgGlyProThrIleApsGlyArgIleApsPro 423
Db      344 -----ATTGTTCTTCGTTTCTTCGCAATGTCCAACATACGACGCTAGAGAAACCC 397
Qy      424 ApsValMetAlaProGlyThrTrpIleLeuSerAlaApsSerSerLeuAlaProApsSer 443
Db      398 GATCTTGTGGTCCCGGGAAGAGGTGTGCTCTTCTCTGCGAAAGCATGACCT 457
Qy      444 SerPheTrpAlaAenHisApsSerGlyAlaTrpMetGlyGlyThrSerMetAlaThr 463
Db      458 AAA-----CAATGTAAGAGTGGGCCAACGCGAGTTTCATCGATGCGACCT 502
Qy      464 ProIleValAla 467
Db      503 GCGGCGGTGCG 514

RESULT 14
LOCUS   CN808705          874 bp      mRNA          linear      EST 27-MAY-2004

```


ORIGIN

Alignment Scores:

Pred. No.:	6,69e-13	Length:	555
Score:	214.00	Matches:	60
Percent Similarity:	50.7%	Conservative:	17
Best Local Similarity:	39.5%	Mismatches:	49
Query Match:	6.6%	Indels:	26
DB:	10	Gaps:	5

US-10-784-870-4 (1-639) x CW803334 (1-555)

```

QY      321  SerGlnAlaPheSerAlaGlyAlaArgIleThrAsnSerTrpGlyAlaValAla 340
      118  TCMAAGACTGCTCCAAAGCAGCAATATTCTTAACACAGTGGGT-----GGC 168
QY      341  GlyAlaTyThrThrAspSerArgAsnValAspAspTyValArgLyAsnAspMetThr 360
      169  GGTCAAGGTCTCAGATGTTGATGGGTTATCAAGCTTGGCAGGTGCGAGGATCATT 228
QY      361  IleLeuPheAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThr 380
      229  CCTGTGTAAGTCAAGGAACTGCTGCTTAACTGCGTACTATGTCATCGCTGGAGAT 288
QY      381  AlaLyAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTy 400
      289  TCGGCTAGTGTGATCAAGTAGGTGCGCCGAT----- 321
QY      401  AlaAspAsnIleAsnHis--ValAlaGlnPheSerSerArgGlyProThrLyAspGly 419
      322  -----ATCAACGACGACCTTGCTCTTTCAAGCAGCAAGCCCGACAGTCCAGGT 372
QY      420  ArgIleLeuPheProAspValMetAlaProGlyThrTyIleLeuSerAlaArgSerSerLeu 439
      373  CTCGAAACCGGATGGCAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
QY      440  AlaProAspSerSerPheTrpAlaAsnHisAspSerTySerAlaTyMetGlyGlyThr 459
      415  -----TCATCTTGCTGACGAC--GACGCTCTTACTGCTTCAATCGGCTCC 462
QY      460  SerMetAlaThrProIleValAlaGlyAsnValAla 471
      463  AGTATGGCTGCACCGCAGTCCGACAGTGCATCGCT 498

```

Search completed: April 8, 2006, 07:51:36
 Job time : 6643 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 8, 2006, 03:16:46 ; Search time 520 Seconds
(without alignments)
2184.350 Million cell updates/sec

Title: US-10-784-870-4

Perfect score: 3267
Sequence: 1 MRKXVFLSVLSAAILSTV.....EVOAYNVPGQNFSLATIV 639

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xip
-Q=/abs/ABSSMB.epool/US10784870/runac.07042006_143420_27544/app_query.fasta_1
-DB=Issued Patents NA -QFMT=fastap -SUFIX=p2n.rml -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10784870.OCGN_1_1_512_@runac.07042006_143420_27544
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMECUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/1.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5.COMB.seq:*
3: /cgn2_6/prodata/1/ina/68.COMB.seq:*
4: /cgn2_6/prodata/1/ina/68.COMB.seq:*
5: /cgn2_6/prodata/1/ina/H.COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCTUS.COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP.COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE.COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3267	100.0	1920	3	US-09-509-814A-3
2	3267	100.0	1920	3	US-09-920-954-3
3	3072	94.0	1923	3	US-09-509-814A-7
4	3072	94.0	1923	3	US-09-920-954-7
5	3066	93.8	1923	3	US-09-509-814A-5
6	3066	93.8	1923	3	US-09-920-954-5
7	2799	85.7	3003	2	US-08-873-479-41
8	477	14.6	1977	3	US-08-894-818B-2
9	477	14.6	1977	3	US-09-445-472-11

10	477	14.6	1977	3	US-10-090-624-11
11	477	14.6	1977	3	US-09-841-553-2
12	443	13.6	1962	3	US-08-894-818B-34
13	443	13.6	1962	3	US-09-445-472-15
14	443	13.6	1962	3	US-10-090-624-15
15	443	13.6	1962	3	US-09-841-553-34
16	443	13.6	9125	3	US-09-902-540-959
17	432	13.2	1977	3	US-08-894-818B-6
18	432	13.2	1977	3	US-09-841-553-6
19	417.5	12.8	1236	3	US-09-445-472-2
20	417.5	12.8	1236	3	US-10-090-624-2
21	417.5	12.8	1566	3	US-08-894-818B-4
22	417.5	12.8	1566	3	US-09-841-553-4
23	362.5	11.1	2539	3	US-09-000-016-3
24	362.5	11.1	2539	3	US-09-514-340-3
25	357.5	10.9	2809	3	US-09-000-016-1
26	357.5	10.9	2809	3	US-09-514-340-1
27	326	10.0	2835	2	US-08-750-532-2
28	326	10.0	4765	2	US-08-750-532-8
29	326	10.0	4765	3	US-08-894-818B-7
30	326	10.0	4765	3	US-09-445-472-5
31	326	10.0	4765	3	US-10-090-624-5
32	326	10.0	4765	3	US-09-841-553-7
33	322	9.9	2532	2	US-07-671-376C-4
34	316.5	9.7	1306	3	US-09-966-921A-1
35	316.5	9.3	1170	3	US-09-902-540-8987
36	303.5	9.3	1170	3	US-08-894-818B-15
37	303.5	9.3	1859	3	US-09-841-553-15
38	303.5	9.3	898	2	US-08-750-532-7
39	292	8.9	898	2	US-08-750-532-7
40	283.5	8.7	564	2	US-08-894-818B-14
41	283.5	8.7	564	3	US-09-841-553-14
42	283.5	8.7	564	3	US-09-841-553-14
43	278.5	8.5	1191	2	US-08-434-255-3
44	278.5	8.5	1191	2	US-08-459-967-3
45	278.5	8.5	1191	2	US-08-460-327-3

ALIGNMENTS

RESULT 1
US-09-509-814A-3
; Sequence 3, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(1920)
; US-09-509-814A-3

Alignment Scores:

Prod. No.: 6,45E-311 Length: 1920
Score: 3267.00 Matches: 639
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-784-870-4 (1-639) x US-09-509-814A-3 (1-1920)

```
QY 1 MetArgLysIleValPheLeuSerValLeuSerAlaAlaIleLeuSerThrVal 20
DB 1 ATGAGAAAGAAAGAGGTTTATCTGTTTATCAGCTGCAGCATTCGTGCAGCTGT 60
QY 21 AlaLeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheIle 40
DB 61 GCATTAAACATCCCTCGCTGGTGAAGCAAGCATTTGATCTGATTTAAAGAAATT 120
QY 41 GlnThrThrAspValSerGlyPheSerGlnArgGlnThrGlyAlaAlaIlePhe 60
DB 121 CAACCAACACGATGTCAGTGGTTTCTCCAAACAGCAACAGTGGCGCTGCATTT 180
QY 61 LeuValGlnSerGluAsnValIleLeuLeuIleGlyLeuLeuIleValIleVal 80
DB 181 CTGGTGAAGTCGAAATGGAATCTTAAAGCATTTGCTTAAAGAACTTGAAACAGTA 240
QY 81 ProAlaAsnAsnIleLeuHisIleValGlnPheAsnGlyProIleLeuGlnGlnThr 100
DB 241 CCGGCAAAATTAATTAACCTCCATATGTCATTAAGTGGCCCATTTTGAAGAAACAAA 300
QY 101 GlnIleLeuGlnIleThrThrGlyAlaIleValIleLeuAspThrIleProAspThrAlaIle 120
DB 301 CAGAACTAGAGCAATCGAGCAAGATTCCTGCATCAGCATCCCTGATTTATGATATTT 360
QY 121 ValGlnIleGlnIleValAspValGlnSerIleValIleValIleValIleVal 140
DB 361 GTCGATATGAGGGGATGTCAGTCAGCAAAAGTCCTGCAATGAAACGCGGAATACGTG 420
QY 141 GlnProIleLeuProIleValIleAspProIleLeuPheThrIleGlyIleAspThr 160
DB 421 GAGCCATCTCTGCCAAATTAACAAATTAAGATCCCACTTTTCAAAAGGCGCATGAGC 480
QY 161 LeuValIleValIleValIleLeuAspThrIleGlnAsnAsnIleValIleValIleVal 180
DB 481 CTGGTGAAGCGTGGCGCTTGAATGAAAGCAAGCAATTAAGAAATGCAATTAAGAGC 540
QY 181 IleGlnIleIleAlaGlnIleValIleAspAspValIleIleThrIleThrAlaIlePro 200
DB 541 ATCGAGGAATCGCTCAGTACGTAGCAACAAATGACGTCATTTATTCGGGCAAGCCT 600
QY 201 GlnIleValIleValIleAspAspValIleValIleValIleValIleValIleVal 220
DB 601 GAATATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 221 SerIleValIleValIleValIleValIleValIleValIleValIleValIleVal 240
DB 661 AGCTACGCTTGTGATGAGCAAGCGCATTTGCGCATTCGCGCATCTGATTTGATTA 720
QY 241 GlyArgAsnAspSerSerMetHisGlnIlePheAspGlyIleValIleThrAlaIleVal 260
DB 721 GGAAGAAAGCAAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 261 LeuGlyArgThrAsnAsnIleAsnAspThrAsnGlnIleValIleValIleValIleVal 280
DB 781 CTGGGTGGAACCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 281 ValLeuGlyAsnGlyAlaIleThrAsnIleValIleValIleValIleValIleValIleVal 300
DB 841 GTATTATGAAATGCGCGCAACGAATTAAGAAATGCGCATTCGCGCATTTGATTTGAA 900
QY 301 SerIleMetAspSerSerIleValIleValIleValIleValIleValIleValIleVal 320
DB 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
```

```
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerThrGlyAlaAlaValAsn 340
DB 961 AGCCAAAGCATTAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 1020
QY 341 GlyAlaIleThrThrAspSerArgAsnValIleAspAspThrValIleValIleValIleVal 360
DB 1021 GGGGCGCTACACAGCATGATTCAGAAATGTCATGATGATGATGATGATGATGATGATGATGAT 1080
QY 361 IleValPheAlaIleGlyAsnGlnIleProAsnGlyIleThrIleSerAlaProGlyThr 380
DB 1081 ATTTCTTTTTCGCGCTGGGATGAAAGCGCAAGCGGATCAGATGCACTTGATACG 1140
QY 381 AlaIleAsnAlaIleThrValIleValIleValIleValIleValIleValIleValIleVal 400
DB 1141 GCTAAACAGCATTAACAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAT 1200
QY 401 AlaAspAsnIleAsnHisIleValIleGlnPheSerSerArgIleProThrIleAspGlyArg 420
DB 1201 GCAGATTAATTAACCACTTGCAGATTCCTTCCCGTGGCCGACAAAGATGGCGCA 1260
QY 421 IleValProAspValMetAlaProGlyThrThrIleLeuSerAlaArgSerSerLeuAla 440
DB 1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 441 ProAspSerSerPheThrAlaAsnHisAspSerIleValIleValIleValIleValIleVal 460
DB 1321 CCGCATCTCTCTTGGGCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 461 MetAlaThrProIleValIleValIleValIleValIleValIleValIleValIleValIleVal 480
DB 1381 ATGGCAACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 481 ArgGlyIleThrProIleProIleLeuLeuValIleValIleValIleValIleValIleVal 500
DB 1441 AGAGCAATCACTCTCAAGCTCTCTCAATTAAGCAAGCTTGAATGATGATGATGATGATGATGAT 1500
QY 501 ValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 520
DB 1501 GTTGAATGGGTTATCCCAACGAAACCAAGATGGGCGGATGATGATGATGATGATGATGATGAT 1560
QY 521 LeuAsnValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 540
DB 1561 TTGAACGTTGCTTATGTAAGCAATCCAGTCCCTATCACTGCAACAAAGCGCATAT 1620
QY 541 ThrPheThrAlaThrAlaGlyIleProIleLeuValIleSerLeuValIleThrAspAlaPro 560
DB 1621 ACCTTTACTGCAACGCGCGGCAAGCATTTGAATCTCTCGTATGATGATGATGATGATGATGAT 1680
QY 561 AlaSerThrThrAlaSerValIleValIleValIleValIleValIleValIleValIleVal 580
DB 1681 GCAGCACTTACTGCTTCTGTAACCTGATCAATGATTTGATTTGATTTGATTTGATTTGATTTGAT 1740
QY 581 AsnGlyThrArgIleValIleValIleValIleValIleValIleValIleValIleValIleVal 600
DB 1741 AACGCAACAAATATGTCGGGATGATCTTCAGACCACTTTGCAATTAACGGAATGCG 1800
QY 601 ArgAsnAsnValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 620
DB 1801 CGCAATTAACGTAAGAAATGATTTATTAATTCGCCCCCAAGTGAACATATTAACATTTGAG 1860
QY 621 ValGlnAlaIleValIleValIleValIleValIleValIleValIleValIleValIleVal 639
DB 1861 GTGCAAGATTAATATGTCGCGTTGCAACCAAACTTCTGTTGGCAATTTGTAAAC 1917
```

RESULT 2

US-09-920-954-3

Sequence 3, Application US/09920954

Patent No. 6759228

GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO

APPLICANT: OKUDA, MITSUYOSHI

APPLICANT: SAKI, KATSUHIISA

APPLICANT: KUBOTA, HIROMI

QY 561 AlaSerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaPro 580
| | | | |
Db 1681 GCAGPACTCTCTCTCTCTGTAACCTGGTCAAGATGATGGATTTGGTCATTACAGACCA 1740
| | | | |
QY 581 AsnGlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTTPAspGly 600
| | | | |
Db 1741 AACGGAACAAGATATGTCGGGAATGACTTCTCAGCACCATTTGACATATACCTGGATGGC 1800
| | | | |
QY 601 ArgAsnAsnValGluLeuAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGlu 620
| | | | |
Db 1801 CGGAATAACCTGAATAATGATTAATTAATTCGCCCAAGTGAACATATATACCATTTGAG 1860
| | | | |
QY 621 ValGlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
| | | | |
Db 1861 GTCCAAAGCATATATATGTCGGTTGACACAAATCTTCTCGTTGGCATTTGTGAAC 1917
| | | | |
RESULT 3
US-09-509-814A-7
/ Sequence 7, Application US/09509814A
/ Patent No. 6376227
/ GENERAL INFORMATION:
/ APPLICANT: TAKAIWA, MIKIO
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SAKKI, KATSUHIKA
/ APPLICANT: KUBOTA, HIROMI
/ APPLICANT: HITOMI, JUN
/ APPLICANT: KAGEYAMA, YASUSHI
/ APPLICANT: SHIKATA, SHITSUMI
/ APPLICANT: NOMURA, MASAFUMI
/ TITLE OF INVENTION: ALKALINE PROTEASE
/ FILE REFERENCE: 0327-0832-0PCT
/ CURRENT APPLICATION NUMBER: US/09/509, 814A
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: PCT/JP98/04528
/ PRIOR FILING DATE: 1998-10-07
/ PRIOR APPLICATION NUMBER: JP 9-274570
/ PRIOR FILING DATE: 1997-06-08
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 7
/ LENGTH: 1923
/ TYPE: DNA
/ ORGANISM: Bacillus sp.
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1923)
US-09-509-814A-7
Alignment Scores:
Pred. No.: 9, 38e-292
Score: 3072.00 Length: 1923
Percent Similarity: 97.0% Matches: 566
Best Local Similarity: 93.4% Conservative: 23
Query Match: 94.0% Mismatches: 19
DB: 3 Gaps: 0
US-10-784-870-4 (1-639) x US-09-509-814A-7 (1-1923)

Db 187 GTGAATCTGAATAATGTGAACCTCCAAAGGTTTGACAGAAGCTTGAACAGTCCG 246
| | | | |
QY 82 AlaAsnAsnValLeuValIleValGlnPheAsnGlyProIleLeuGlnThrTyrGln 101
| | | | |
Db 247 GCAAATTAATAACTCCATATATATCCAAATGACCAATTTTGAAGAAACAAACAG 306
| | | | |
QY 102 LysLeuGlnThrThrGlyValAlaValIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
| | | | |
Db 307 CAGCTGGAAAAACAGGGGCAAGATTTCTGACTACATCTGATTAATGCTTACATTGTC 366
| | | | |
QY 122 GlnTyrGlnGlyAspValGlnSerLysValArgSerIleGluIleValGluSerValGlu 141
| | | | |
Db 367 GAGTATGAGGGCGAATTTAGTCAGAACACACCATTTGACAGCTGGAAATCCGTGAG 426
| | | | |
QY 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThrLeu 161
| | | | |
Db 427 CCTTATTTCCCATATACAGAAATAGATCCCGCTTTTCACAAAGGGGCAACAGACTT 486
| | | | |
QY 162 ValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnValGluValGlnLeuArgGlyIle 181
| | | | |
Db 487 GTAAAGACGTGGCGCTTGATACAAAGCAAAATTAAGAGTGCATTAAGAGGCAATC 546
| | | | |
QY 182 GlnGlnIleAlaGlnTyrValAlaSerAsnAspValIleTyrIleThrAlaLysProGln 201
| | | | |
Db 547 GAACAAATCGCAAAATTCGCAATTAAGCAATGATGTGCTATATATACGGCAAAAGCTGAG 606
| | | | |
QY 202 TyrIleValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSer 221
| | | | |
Db 607 TTTAAGGTGATAAAGATGTTCGCCGTGAATTTGCAAAAGTGTGCTCAGAGCAGC 666
| | | | |
QY 222 TyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
| | | | |
Db 667 TACGGGTGTATGACAAAGACAGATCTGAGCGGTGGCAATACAGGCTTGATACAGGT 726
| | | | |
QY 242 ArgAsnAspSerSerMetHisGlnAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeu 261
| | | | |
Db 721 CGCAATGACAGTTTCATGATGATGAACTTCGGCGGAAATTAATCGCATTAATGATGATG 786
| | | | |
QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerVal 281
| | | | |
Db 787 GACCGAGAGAAATTAAGCCAAATGATGATGATGATGATGATGATGATGATGATGATG 846
| | | | |
QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
| | | | |
Db 847 TTAGAAACGGCTCCATTAATTAAGAAATGGCGCTTCAGCGAAATCTTGCTTCCAAATCT 906
| | | | |
QY 302 IleMetAspSerSerGlyLysGlyLysLeuProSerAsnLeuGlnThrLeuPheSer 321
| | | | |
Db 907 ATCATGATGATGCGGTGGGGGACTTGAGAGACTACCTTGCAATCTGCAAACTTATTCAGC 966
| | | | |
QY 322 GlnAlaPheSerAlaGlyValArgIleHisThrAsnSerTTPGlyAlaAlaValAsnGly 341
| | | | |
Db 967 CAAAGCATTCAGTGTGGGCGCAAGATTCATCAAACTCTGGGGGAGCAGTGAATGCG 1026
| | | | |
QY 342 AlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThrIle 361
| | | | |
Db 1027 GCTTACAAACAGATTCAGAAATGTGATGCTATGTGCGCAAAATGAATATGACATC 1086
| | | | |
QY 362 LeuPheAlaAlaGlyAsnGlnArgProAsnGlyTyrThrIleSerAlaProGlyThrAla 381
| | | | |
Db 1087 CTTTTCGTGGCGGAAATGAAGAACCGAACCGGGAACATCAAGTGCACAGGCACT 1146
| | | | |
QY 382 LysAsnAlaIleThrValGlyAlaThrGluLeuLeuAspProSerPheGlySerTyrAla 401
| | | | |
Db 1147 AAAAATGAATAACAGTGGAGCTACGGAACAACTCCCAAGCTTTGGGTCTTATGCG 1206
| | | | |
QY 402 AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIle 421
| | | | |
Db 1207 GACAAATATCAACCAATGTGACAGTTCCTTCACTGAGCGCAAAAGAAATGAGAGATC 1266
| | | | |
QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
| | | | |
Db 1267 AAACCGAGATGTACAGCCGGGAACGTTCAATATCATAGCAAGATCTTCTCTTGACAGC 1326
| | | | |


```
QY 442 AppSerPheThraAlaAmh1aAppSerIyTrAlaTrMetGlyGlyThrSerMet 461
Db 1327 GATTCCTCTCTTCGGGGAACCATGACATGATTAATATCATATGATGGTGAACCTCATG 1386
QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlnHisPheValIlyAsnArg 481
Db 1387 GCTAACCCGATCGTGTGTGAACGTGGCAAGCTTGCTGATGACATTTTGGAAAAACGA 1446
QY 482 GlyIleThrProIlyPProSerLeuLeuValAlaLeuIleAlaGlyAlaAlaPheVal 501
Db 1447 GGCATCAACCAAGCTTCTCTATTAAGCGCACGTGATGCGGTGACGCTGACATC 1506
QY 502 GlyLeuGlyTrProAsnGlyAsnGlnGlyTrpGlyValIleThrAlaProAsn 521
Db 1507 GGCCTTGCTACCGAAGCGTACCAAGATGGGGAAGATGACATTTGAATATCCCTG 1566
QY 522 AsnValAlaTrValAlaGlnGlySerSerAlaLeuSerThrSerGlnIlyAlaThrTrpThr 541
Db 1567 AACGTGGCTATGATGAACGATCCAGTTCCTATCCACGACCAAAAGGACGTACTCG 1626
QY 542 PheThrAlaThrAlaGlyIlyProLeuIlyIleSerLeuValIlySerAspAlaProAla 561
Db 1627 TTTACTGCTACTCGCGCAAGCCTTGAATAATCTCCCTGGTATGCTGATGCCCTGCG 1686
QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCAACAACGCTTCCTCCGACGCTGTGCAATGATCTGACCTGTGCATTAACCTCCAAAT 1746
QY 582 GlyThrArgTrValAlaGlyAsnAspPheSerAlaProPheAspAsnAntTPAspGlyArg 601
Db 1747 GGCACACAGTATATAGAAATGACTTACTTCCCATACATATGATATGCGATGCGCGC 1806
QY 602 AsnAsnValGlnAsnValPheIleAsnSerProGlnSerGlyThrTrpIleGlnVal 621
Db 1807 AATTAAGTGAATAATGATTTATTAATGACACCAAAAGCGGACGTATCAATTGAAGTA 1866
QY 622 GlnAlaTrpAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGATTAATACGATCCGATGGTGAACACAGAACTTCCTGTTGCAATTTGAAT 1920

RESULT 4
US-09-920-954-7
; Sequence 7, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
```

```
US-09-920-954-7
Alignment Scores:
Pred. No.: 9,38e-292 Length: 1923
Score: 3072.00 Matches: 596
Percent Similarity: 97.0% Conservative: 23
Best Local Similarity: 93.4% Mismatches: 19
Query Match: 94.0% Indels: 0
DB: 3 Gaps: 0

US-10-784-870-4 (1-639) x US-09-920-954-7 (1-1923)
QY 2 ArgIlyLeuIleValAlaPheLeuSerValIleuSerAlaAlaIleuSerThrValAla 21
Db 7 AAGAGGAAAAAGGATGTTTATCTGTTTATCACTGCAAGCGATTTGTGACTGTGCG 66
QY 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheIlyGlnIleGln 41
Db 67 TTAAGTATCATCTCAGGTGGTGCAGAGGAATTTTGATCTGATTTCAAGGAATTCAG 126
QY 42 ThrThrThrAspValSerGlyPheSerIlyGlnArgGlnThrGlyAlaAlaIlePheLeu 61
Db 127 ACAACAACGATGCTAAGGTTTCTCCAAAGCGGCGACAGTGTGCTGCTTTCTG 186
QY 62 ValGluSerGlnAsnValIlySerLeuLeuGlyIleuLeuIlyValIleuGlnIleuPro 81
Db 187 GTGGAATCTGAATAATGTAACCTCCCAAGGTTTCAAGAGAGCTTGAAACGATCCG 246
QY 82 AlaAsnAsnIlySerLeuHisIleValAlaGlnPheAsnGlyProIleuGlnIleuThrIlyGln 101
Db 247 GCAATATATATATCTCATATTTATTCATTCATTCATTCATTCATTCATTCATTCATTC 306
QY 102 IlyLeuGlnIleuThrThrGlyAlaIlyIleuAspTrpIleProAspTrpAlaTrpIleVal 121
Db 307 CAGCTGAAAAAACAAGGCGCAAGATTCGATCAATCACTGATTAATGCTTCAATGTC 366
QY 122 GlnTrpGlnGlyIlyAspValGlnSerIlyValArgSerIleGlnHisValGlnIleuGln 141
Db 367 GAGTATGAGGCGCATTTAATGATGCAACAGCAACCAATTAAGCACTTGATGATCGTGGAG 426
QY 142 ProTrpLeuProIlySerIlyIleAspProGlnIleuPheThrIlyGlyAlaSerThrLeu 161
Db 427 CCTATTTCCTCATATACAGAAATGATCCCGACTTTTCAAAAGGCGATCAGAGCTT 486
QY 162 ValIlyAlaLeuAlaLeuAspThrIlyGlnAsnAsnIlyGlnValGlnLeuArgGlyIle 181
Db 487 GTAAAAAGCAGTGGCGCTGATACAAAGCAGAAAAAATTAAGAGGTGCAATTAAGAGCATC 546
QY 182 GlnGlnIleAlaGlnTrpValAlaSerAsnAspValHisTrpIleThrAlaIlyProGln 201
Db 547 GAACAATGCAACAAATTCGCAATTAACCAATGATGCTATATTAATTCGGCAAGCCTGAG 606
QY 202 TrpIlyValIleAsnAspValAlaArgGlyIleValIlyAlaAspValAlaGlnSerSer 221
Db 607 TATAGGTATGATGATATGTTGCGGTGATTTGTCAAAAGCGGATGTGCTCAGAGCAGC 666
QY 222 TrpGlyLeuTrpGlyGlnIlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
Db 667 TACGGGTGTATGACAAAGCAAGATCGTAGCGTTGCCGATCAAGGCTTGATACAGGT 726
QY 242 ArgAsnAspSerSerMetHisGlnAlaPheAspGlyIlyIleThrAlaLeuTrpAlaLeu 261
Db 727 CGCAATGACAGTTCATGATGATGAGCTTCGCGGGAATAATTAATGCAATTAATGCAATG 786
QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerVal 281
Db 787 GAGCGACAAATTAATCCAAATGATGATGATGATGATGATGATGATGATGATGATGATG 846
QY 282 LeuGlyAsnGlyAlaThrAsnIlyGlyMetAlaProGlnAlaAsnLeuValIleGlnSer 301
Db 847 TTAGAAACGGCTCCACTAATTAAGGAATGGCGCTCAGGCGAATCTTCAATCTT 906
QY 302 IleMetAspSerSerGlyIlyLeuGlyIlyLeuProSerAsnLeuGlnThrLeuPheSer 321
```

Db 907 ATCATGATGATGCGGTGGGAGCTTGAGGACCTTCGAAATCTGCAAACTTATTACAG 966
Qy 322 GlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTrpGlyAlaValaIasnGly 341
Db 967 CAGGCAATACAGTCTGGTGCAGAAATTCATACAACTCCGGGAGACAGAGAAATGGG 1026
Qy 342 AlaTyrThrThraPseSerArgenValaAspAspTyrValaArgIleAsnAspMetTrile 361
Db 1027 GCTTACACAAAGATTCGAAATGCTGAGTACTATGTGGCAAAATGATATGACGATC 1086
Qy 362 LeuPheAlaIaGlyAsnGlyArgProAsnGlyGlyThrIleSerAlaProGlyThrAla 381
Db 1087 CTTTGGCTGGCCGGAGATGAGACCGAAGCGGGAACCATGTCACACAGGACAGCT 1146
Qy 382 LysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
Db 1147 AAAAAATGCAATACAGTCTGGAGCTACGAAAMACCTCCGCAAGCTTGGCTTATGGC 1206
Qy 402 AspAsnIleAsnHisValaIaGlnPheSerSerArgGlyProThrLysAspGlyArgIle 421
Db 1207 GACAAATACACAGATTCGACAGCTTCTTTCACGTGACCGAAGATGACGAGATC 1266
Qy 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
Db 1267 AAACCGAGATGTCATGGCACCGGGAACGTTCACTATCAGCAAGATCTTCTTGACACCG 1326
Qy 442 AspSerSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMetCylGlyLysSerMet 461
Db 1327 GATTCCTCTCTGGGCGAACCATGACAGTAAATATGCAATACATGGGTGGAACGTCCAGT 1386
Qy 462 AlaThrProIleValaIaGlyAsnValaIaGlnLeuArgGluHisPheValIysAsnArg 481
Db 1387 GCTACACCAATCGTCTCTGGAACGTGACAGCTTCGTGAGCATTTTGTGAAAAACAA 1446
Qy 482 GlyIleThrProLysProSerLeuLeuValaIaIaLeuIleAlaGlyAlaIaAspVal 501
Db 1447 GGCATACACCAAGGCTTCTCATTTAAAGCGGACATGATGCGGTGACGTGACATC 1506
Qy 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValaThrLeuAspLysSerLeu 521
Db 1507 GGCCTTGCTACCCGAAACGCTAAACAGATGGGAGAGTGCATGTGATTAATCCCTG 1566
Qy 522 AsnValaIaTyrValaIasnGlySerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
Db 1567 AACGTGCTATGTGAACAGATCCAGTCTCTATCCACAGCCAAACAGACGTACTCG 1626
Qy 542 PheThrAlaIaThrAlaGlyLysProLeuLysIleSerLeuValaTyrSerAspAlaProAla 561
Db 1627 TTTACTGCTACGCGGCAAGGCTTGAATCTCCCTGTATGCTGATGCCCTCGG 1686
Qy 562 SerThrThrAlaSerValaThrLeuValaAsnAspLeuAspLeuValaIleThrAlaProAsn 581
Db 1687 AACCAACAGCTTCGCTAACGCTGTCAATGATGAGACCTTGCTCATCCCTCCAAAT 1746
Qy 582 GlyThrArgTyrValaGlyAsnAspPheSerAlaProPheAspAsnAsnTrpAspGlyArg 601
Db 1747 GGCACACAGATATGAGAAATGATCTTACTCGCCATACAAATGATTAACGTGAGATGCC 1806
Qy 602 AsnAsnValaGluAsnValaPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
Db 1807 AATAACGTAGAAATATGATTTATTAATGACACCAAGGCGGAGCATGATCAATTGAAGTA 1866
Qy 622 GlnAlaTyrAsnValaProValaGlyProGlnAsnPheSerLeuAlaIleValaAsn 639
Db 1867 CAGGCTTATACGATCGGTGACCAAGAACTTCTCGTTGCAATGTGAAAT 1920

RESULT 5

US-09-509-814A-5
; Sequence 5, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIIWA, MIKIO

APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIRATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (1923)
US-09-509-814A-5
Alignment Scores:
Pred. No.: 3,656-291 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.94 Conservative: 23
Best Local Similarity: 93.34 Mismatches: 20
Query Match: 93.84 Indels: 0
Gaps: 0
US-10-784-870-4 (1-639) x US-09-509-814A-5 (1-1923)
Qy 2 ArgLysLysValaPheLeuSerValaLeuSerAlaIaIaIleLeuSerThrValaIa 21
Db 7 AAGAGGAAAGAGTCTTTTATCTGTTTATCATGCTGACGCAATTTGTGCACTGTCG 66
Qy 22 LeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
Db 67 TTAAGTAAATCCATCGCAGGTGTCGCAAGAAATTTGATCTGATTTCAAGAAATTCAG 126
Qy 42 ThrThrThrAspValaSerGlyPheSerLysGlnArgGlnThrValaIaIaPheLeu 61
Db 127 ACAACACTGATGCTAAAGGTTTCTCCAAAGAGGCGAGCTGCTGCTGCTTTCTG 186
Qy 62 ValGluSerGluAsnValaLysLeuLeuLysGlyLeuLeuLysLysGlyuThrValaPro 81
Db 187 GTGGAATCTGAAGATGTAACCTCCCAAGGTTTGCAGAAAGAACTTGAAACAGTCCCG 246
Qy 82 AlaAsnAsnLysLeuHisIleValaGlnPheAsnGlyProIleLeuGluGluThrLysGln 101
Db 247 GCAATATATTAATCTCATATATTCATCAATGACCAATTTAAGAAACAAACAG 306
Qy 102 LysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
Db 307 CAGCTGAAAAAAGCGGGCAAGATCTTCACATCAATCACTGATTAATGCTTACATGTC 366
Qy 122 GluTyrGluGlyAspValaIaGlnSerLysValaLysSerIleGluHisValaGluSerVala 141
Db 367 GAGTATGAGGCGATGTTTAATCAGCAACAGACATGAGCAGTGGAAATCCGTGAG 426
Qy 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaIaSerThrLeu 161
Db 427 CTTATTTGCGATATACAGATAGATCCACGCTTTTCAAAAAGGGGCAATCAGAGCTT 486
Qy 162 ValLysAlaLeuAlaIaLeuAspThrLysGlnAsnAsnLysGlyuValaGlnLeuArgIle 181
Db 487 GTAAAAAGAGTGGGCTTGATACAAAGCAAAAAATAAAGAGGTGCATTTAAGAGGATC 546
Qy 182 GluGluIleAlaGlnTyrValaIaIaSerAsnAspValaHisTyrIleThrAlaLysProGlu 201

```

Db      547 GAAACAATCCGACATTCGGCAATAGCAATGATGTCTATATATATAGCGGAAAGCCCTGAG 606
      202 TTTTCTGCTACTGCGCGCAAGCCCTTGAATAATCTCCCTGTATGTCGATGCCCTGCG 1686
Qy      562 SerThrThraLaseValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
      1687 AGCACAACCTGCTCCGTAACGCTTGCAAGATCTGCAACCTTGTCAATTACCGCTCCAAAT 1746
Db      607 TATTAAGGTGATGATGATGTTGGCCGTCGGAATGTCTCAAGCGGATGTGGCTCAGACGACG 666
Qy      222 TTTGTYLeuTYrGlyGlnGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
      667 TACGGGTGTATGACAAAGACAGATGTAGCGGTCCGATACAGGAGCTTGATACAGGT 726
Qy      242 ArgAsnAspSerSerMetHisGluAlaPheArgGlyLeuIleThrAlaLeuValAlaLeu 261
      727 GCGATACACAGTTCGATGCAATGAAGCCTTCGCGGAAATTAATCTGATTAATGACTTG 786
Qy      262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyYhiGlyThrHisValAlaGlySerVal 281
      787 GGAACGACGAATATATGCAATGATACGAATGTCTATGTCATGTCATGTCGCTGCTCCGTA 846
Qy      282 LeuGlyAsnGlyAlaThrAsnIleGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
      847 TTAGGAAACGCGCTCCACTAATTAAGAAATGCGCTCAGGCGAATCTATGCTTCCAAATCT 906
Qy      302 IleMetAspSerSerGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSer 321
      907 ATCATGATGATGCGGTGGGAGCTTGAGAGACTTCTGCAATCTGCAAACTTATTTGACG 966
Qy      322 GlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTrpGlyAlaAlaValAsnGly 341
      967 CAAGCATACAGTGTGTGTGCAAGATTCATACAACTCTCGGGAGCAGCAGTGAATGCGG 1026
Qy      342 AlaIleThrThrAspSerAlaAsnValAspAspTyrValArgIleAsnAspMetThrIle 361
      1027 GCTTACACACAGATTCACAAATGTGATGATCTATGTGGCAAAATGATATGACATCT 1086
Qy      362 LeuPheAlaIleGlyAsnGlyIleArgProAsnGlyGlyThrIleSerAlaProGlyThrAla 381
      1087 CTTTTCGCTCCGCGGAATGAAGACCGAAGCGCGGAACCATCTACTGACACAGGACACGCT 1146
Db      1147 AAAAATGCAATTAACAGTCGAGCTACGGAACCTCCGCCCAAGCTTTGGGTCTTATGCG 1206
Qy      402 AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrIleAspGlyValArgIle 421
      1207 GACAAATATCAACATGTGGCAGAGTCTCTTCACTGTGACGCAAGGATGAGCGCATCT 1266
Qy      422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
      1267 AAACCGGATGTCAATGGACCGGAAACGTTCACTACTACACAGATCTTCTTGGACCGC 1326
Qy      442 AspSerSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMetGlyIleThrSerMet 461
      1327 GATTCCTCTCTTGGGGAACATGACAGTAATATGCAATACATGGGTGAAAGCTCATG 1386
Qy      462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlyIleHisPheValIleAsnArg 481
      1387 GCTAACCCGATCGTGTGGAACGTTGCAACGCTTGTGAGCATTTTGTGAATAACAGA 1446
Qy      482 GlyIleThrProIleAspSerLeuLeuValAlaIleLeuIleAlaGlyAlaAlaAspVal 501
      1447 GGCATTCACACAAAGCTTCTCTATTAAGCGGACATGATGCGCGGTGACGTCGATC 1506
Qy      502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTrpGlyValThrLeuAspLysSerLeu 521
      1507 GGCCTTGGCTACCGGAACGTTAACCAAGATGGGAGATGAGTGAATGAGTAATATCCCTG 1566
Qy      522 AsnValAlaTyrValAsnGlyLysSerSerAlaLeuSerThrSerGlyIleValAlaThrTyrThr 541
      1567 AAGGTGCTTATGTGAACGAGTCCAGTCTCTATCCACAGCCAAAGGAGGACGTAATCTG 1626
Qy      542 PheThrAlaThrAlaGlyLysProLeuValIleSerLeuValTyrSerAspAlaProAla 561

```

```

Db      1627 TTTACTGCTACTGCGCGCAAGCCCTTGAATAATCTCCCTGTATGTCGATGCCCTGCG 1686
Qy      562 SerThrThraLaseValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
      1687 AGCACAACCTGCTCCGTAACGCTTGCAAGATCTGCAACCTTGTCAATTACCGCTCCAAAT 1746
Db      607 TATTAAGGTGATGATGATGTTGGCCGTCGGAATGTCTCAAGCGGATGTGGCTCAGACGACG 666
Qy      582 GlyThrArgThrValGlyAsnAspPheSerAlaProPheAspAsnMetTrpAspGlyArg 601
      1747 GGCACACAGTATGAGAAATGATGATTTACTTCCGCAATCAATATTAATCTGGATGCGCG 1806
Qy      602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrIleGluVal 621
      1807 AATTAAGTAAATAATGATTTATTAATGACCAAGGAGGAGGATTAATCAATTAAGGTA 1866
Qy      622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
      1867 CAAGCTTATACGTAACGTAACGCTTGACACAGACCACTCTCGTTGCAATTTGAAAT 1920
Db      1867 CAAGCTTATACGTAACGTAACGCTTGACACAGACCACTCTCGTTGCAATTTGAAAT 1920

RESULT 6
US-09-920-954-5
; Sequence 5, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKETI, KATSUHIKA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-09-920-954-5

Alignment Scores:
Pred. No.: 3,65e-291 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 3 Gaps: 0

US-10-784-870-4 (1-639) x US-09-920-954-5 (1-1923)
Qy      2 ArgIleLysIleValPheLeuSerValLeuSerAlaAlaIleLeuSerThrValAla 21
      7 AAGAAAGAAAGAGTGTATTTATCTGTTTATGCTGACGACGATTTTGTGACTGTGGCG 66
Db      22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
      67 TTAAGTAATCCATCTCAGGTGTGCAAGAAATTTGATCTGATTTCAAGGAATTCAG 126
Qy      42 ThrThrThraAspValSerGlyPheSerLysGlnArgIleThrGlyAlaAlaPheLeu 61
      127 ACAACAATGATGCTTAAGGTTTCTCAAGACAGGCGCAACTGTGCTGCTGCTTTCTG 186

```

QY 62 ValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLeuGluThrValPro 81
 DB 187 GTGGATCTGAAATGTAATGAACTCCCAAAAGTTTCAGAAAGAGCTTGAACAGTCCG 246
 QY 82 AlaAsnAsnLysLeuLysIleValGlnPheAsnGlyProIleLeuGluGluThrLysGln 101
 DB 247 GCAAAATATTAATCTCCATATTAATCAATTCATGAGCAATTTTACAGAAACCAAAACG 306
 QY 102 LysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
 DB 307 CAGCTGGAAAAACAGGGCCAAAGATTCTGACATACCTGATATCTTATCTTACATTTGTC 366
 QY 122 GluTyrGluGlyAspValGlnSerLysValArgSerIleGluIleValGluSerValGlu 141
 DB 367 GAGTATGAGGGCGATTTAAGTCAGCAACAGCAATGAGCAGCTGAAATCCGGAGG 426
 QY 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyLaseThrLeu 161
 DB 427 CCTTATTTCCGATATACAGAAATAGATCCCGAGCTTTTCACAAAAGGGCGATCAGAGCTT 486
 QY 162 ValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgGlyIle 181
 DB 487 GTAAAAGCAGTGGCGCTGATACAAAGCAAAATTAAGAGGTGCAATTAAGAGGCATC 546
 QY 182 GluGluIleAlaGlnTyrValAlaSerAsnAspValIleTyrIleThrAlaLysProGlu 201
 DB 547 GACCAATTCGCAATTCGCAATTCGCAATTCGCTATATATTTACGCAAAAGCTTGAG 606
 QY 202 TyrLysValIleAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSer 221
 DB 607 TATAAGTGATGATGATGATGTTGGCGGTGAATTTGTCAAAAGCGGATGGCTCAGAGCAC 666
 QY 222 TyrGlyLeuTyrGlyGlnGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
 DB 667 TACGGGTGTATGACAGACAGACAGATCGTAGCGGTGGCATACAGGGCTTATATCAGGT 726
 QY 242 ArgAsnAspSerSerMetIleGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeu 261
 DB 727 CGCATGACAGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
 QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyIleGlyThrIleValAlaGlySerVal 281
 DB 787 GAGCGGACGAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 846
 QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
 DB 847 TTAGGAAACGGCTCCCATATTAAGGAATGGCGCTCAGCGCAATCTTACTTCCATCT 906
 QY 302 IleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSer 321
 DB 907 ATCATGATATGCGGTGGGAGCTTGGAGGACTTACCTTCATATCTGCAAACTTATATCCG 966
 QY 322 GlnAlaPheSerAlaGlyAlaArgIleIleIleIleIleIleIleIleIleIleIleIle 341
 DB 967 CAAAGCATACAGTGGTGGTCCAGAAATTCATCAAACTCTGGGAGGACGACATGATGG 1026
 QY 342 AlaTyrThrThrAspSerArgAsnValAspAspTyrValAlaLysLysAsnAspMetThrIle 361
 DB 1027 GCTTACACAAAGATTCAGAAATGGATGATGATGATGATGATGATGATGATGATGATG 1086
 QY 362 LeuPheAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAla 381
 DB 1087 CTTTTCGCTGGCGGAAATGAAGACCGAAAGCGGAAACATACATGAGCACAGGCAAGCT 1146
 QY 382 LysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
 DB 1147 AAAAATGCAATACAGTGGAGCTACGAAACCTCCGCAAGCTTGGCTTATGGC 1206
 QY 402 AspAsnIleAsnIleValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIle 421
 DB 1207 GACATATATCAACATGTCGACAGATTTCTTCACTGACGAGCACCAAAAGATGAGAGATC 1266

QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
 DB 1267 AAACGGATGTATGCGACCGGAACTTTCATATCATGCAAAATCTTCTTTCACCG 1326
 QY 442 AspSerSerPheTyrAlaAsnIleAspSerLysTyrAlaTyrMetGlyGlyThrSerMet 461
 DB 1327 GATTCCTCTCTGGCGGCAACCATGACAGTAAATATGATATCATGGGTGAAAGCTCATG 1386
 QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluIlePheValLysAsnArg 481
 DB 1387 GCTACACGATCTGTGCGAAACGTGCACAGCTTCGTGAGCATTTGTGAAAAACGA 1446
 QY 482 GlyIleThrProLysProSerLeuLeuLysAlaIleValIleAlaGlyAlaAlaAspVal 501
 DB 1447 GGCATCACACCAAGCTTCTTATTAAGGCGCATGATTTGGCGGTGACGTACATC 1506
 QY 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeu 521
 DB 1507 GGCCTGGCTACCGAACGTAAACAGATGAGGAGCAGATGACATTTGATTAATCCCTG 1566
 QY 522 AsnValAlaTyrValAsnGluSerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
 DB 1567 AACGTTCCTATGTGAAACAGATCCAGTCTCTATTCACACGCAAAAGCGACGTACCG 1626
 QY 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAla 561
 DB 1627 TTTACTGCTACCTGGCGCAAGCTTTGAAATCTCCCTGGATGATCTGATGCCCTGCG 1686
 QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
 DB 1687 AGCACAACTGCTTCGTAACGCTTGCAATGATCGACCTTGTCATTAACGCTCAAT 1746
 QY 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGlyArg 601
 DB 1747 GGCACACGATGTGAAAGATGATCTTTCGCTCATCAATGATGATGATGATGATGATG 1806
 QY 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
 DB 1807 AATACGTAGAAATGTATTTATTTATGACACCAACAGGAGCTATACATTAAGGTA 1866
 QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
 DB 1867 CAGGCTTAATACGTACCGGTTGACACACAGACCTTCTGTTGGCAATGTGAAT 1920

RESULT 7
 US-08-873-479-41
 ; Sequence 41, Application US/08873479
 ; Patent No. 5891701
 ; GENERAL INFORMATION:
 ; APPLICANT: Sloma, Alan
 ; TITLE OF INVENTION: Nucleic Acid Encoding A Polypeptide
 ; TITLE OF INVENTION: Having Protease Activity
 ; NUMBER OF SEQUENCES: 57
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: No. 58917010 No. 5891701disk of No. 5891701th America
 ; STREET: 405 Lexington Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10174
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/873,479
 ; FILING DATE: 12-JUN-1997
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Agilis, Cheryl H
 ; REGISTRATION NUMBER: 34,086

REFERENCE/DOCKET NUMBER: 5251.000-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-878-9655
 TELEX:
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3003 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULAR TYPE: Genomic DNA
 US-08-873-479-41

Alignment Scores:

Pred. No.:	1.27e-264	Length:	3003
Score:	2799.00	Matches:	541
Percent Similarity:	92.4%	Conservative:	52
Best Local Similarity:	84.3%	Mismatches:	45
Query Match:	85.7%	Indels:	4
DB:	2	Gaps:	2

US-10-784-870-4 (1-639) x US-08-873-479-41 (1-3003)

QY 1 MetArglyS-----LysLysValPheLeuSerValLeuSerAlaAlaIleLeu 17
 DB 846 ATGAGAAAGAAAGATCGAAGAGGGTTTATCCCTTTATCATGCTGCACTATG 905
 QY 18 SerThrValAlaLeuAenAenProSerAlaGlyAspAlaArgThrPheAspLeuAspPhe 37
 DB 906 TCTTCGTCCTTAAACAGACCTCTTACTATGCGGCGAACAATTTGAATGCACTTT 965
 QY 38 LysGlyIleGlnThrThrThrAspValSerGlyPheSerLysGlnAthrGlnThrGlyAla 57
 DB 966 AAGGGGATAGACACTTACGCTAGAGAGGCTGCCCAAGCAAGCAAGCAAGCAAGCAAG 1025
 QY 58 AlaAlaPheLeuValGluSerGluAenValLysLeuLysGlyLeuLeuLysLysLeu 77
 DB 1026 GCATCTTTCTTTGAACTCTGAAATGTAATCCCAAGACTATTCAAAGAAACTA 1085
 QY 78 GluThrValProAlaAenAenLysLeuAlaIleValGlnPheAenGlyProIleLeuGlu 97
 DB 1086 GAAGTAGTTCAGCGGATTAACAAGCTATATGCTTCAATTTGACGACCTATTTAGAG 1145
 QY 98 GluThrLysGlnLysLeuGluThrThrGlyAlaLysIleLeuAspLysIleProAspLys 117
 DB 1146 GAAACGCACTTCAACTAGAGAGACCGGAGCGCAAAATCTCGATTACATACCAAGATTAC 1205
 QY 118 AlaThrIleValGluThrGluGluLysAspValGlnSerLysValArgSerIleGluIleVal 137
 DB 1206 GCTTAATATGTCGAAATAGATGAGGAGTAAAGCGCTAACTAACGCAATTCGCAATTG 1265
 QY 138 GluSerValGluProLysLeuProLysLysIleAspProGlnLeuPheThrLysGly 157
 DB 1266 GAATCGGTGACCATATTTACCTTATATTAATGACCCGCAATTTATTTCCAGAGCA 1325
 QY 158 AlaSerThrLeuValLysValAlaLeuAlaLeuAspThrLysGlnAenAenLysGluValGln 177
 DB 1326 GCTTCGAATTTAGTAGAGACAGTAGCTTGAATTAAGCAAGCAAGCAAGCAAGCAAGT 1385
 QY 178 LeuArgGlyIleGluGluIleAlaGlnThrValAlaSerAenAspValIleSerIleLeuThr 197
 DB 1386 TTAAAGGATTTGAAACAAATTCGCCAATACGCAACAATATGATGATTAATACGTAACC 1445
 QY 198 AlaLysProGluThrLysValLysAenAspValAlaArgGlyIleValLysValAspVal 217
 DB 1446 CCAAGCGCTGAATACGAAGTTTGAATGACGTGGCCCGCTGCACTTGGAAAGCAGACGTC 1505
 QY 218 AlaGlnSerSerLysLysLeuThrGlyGlnGlyGlnIleValAlaValAlaAspThrGly 237
 DB 1506 GCAACAAATATACCTTGCTTATATGACACAGACAGATTGTACAGATTGCTGATACGCG 1565
 QY 238 LeuAspThrGluArgAenAspSerSerMetIleGluAlaPheArgGlyLysIleThrAla 257

DB 1566 CTGATATCAGAGAAATATGACAGTTCATGATGAAACATTCGCCGCTAATTAACCGCA 1625
 QY 258 LeuThrAlaLeuGluArgThrAenAenAlaAenAspThrAenGlyIleGlyThrIleVal 277
 DB 1626 CTATATGCACTGGGCGAAGCAATTAACGCAATGATCCAATATGACATGGAACCAATGTT 1685
 QY 278 AlaGlySerValLeuGluAenGlyAlaThrAenLysGlyMetAlaProGlnAlaAenLeu 297
 DB 1686 GCTGGATCTGTGTAGGAAT---GCTACAAATTAAGGATGGACCGCAAGCAACCACTA 1742
 QY 298 ValPheGlnSerIleMetAspSerSerGlyLysGlyLysIleLeuProSerAenLeuGln 317
 DB 1743 GCTTTCAATCTATATGATGATAGTGGTGGAGGCTGGAGAGACTTACTCTATCTACAA 1802
 QY 318 ThrLeuPheSerGlnAlaPheSerAlaGlyAlaArgIleIleThrAenSerTrpGlyAla 337
 DB 1803 ACATTAATTCAGTCACAGCATATAGTGTGAGCGCAAGAAATTCATACGAATTCATGGGGGCT 1862
 QY 338 AlaValAenGluValArgThrThrAspSerArgAsnValAspAspLysValArgLysAsn 357
 DB 1863 CCAGTAAACGCTGCTATACGACAGACTCTCGAAATGTTGATATTAATGAGCAAAAT 1922
 QY 358 AspMetThrIleLeuPheAlaIleGlyAsnGluArgProAenGlyThrIleSerAla 377
 DB 1923 GATATGACATTTCTTTTGGCGCGAATAGAGGACCAAGTACGCTGATCAATCAGTGCA 1982
 QY 378 ProGlyThrAlaLysAenAlaIleThrValGlyAlaThrGluAenLeuArgProSerPhe 397
 DB 1983 CCAAGAAACAGCAAAATATGCAATTAAGTTAGGGGCAACGAAACCTAACCTCAAGCTTC 2042
 QY 398 GlySerThrAlaAspAenIleAenIleValAlaGlnPheSerSerArgLysProThrLys 417
 DB 2043 GGAATCTTAAGCGATATATTAACAATGTTGCTCAATTCCTTCAACAGAGCTCTACTAGA 2102
 QY 418 AspGlyArgIleLysProAspValMetAlaProGlyThrThrIleLeuSerAlaArgSer 437
 DB 2103 GATGACGATTTAAGCGGACGATCAATGCAACGAGTACGATATTCCTGCTGACAGTCA 2162
 QY 438 SerLeuAlaProAspSerSerPheTrpAlaAenIleAspSerLysTrpAlaLysMetGly 457
 DB 2163 TCATTTAGCTCAATCTCTCATTCCTGGGCAACAGATAGTAATATGCTTACATGCGT 2222
 QY 458 GlyThrSerMetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluIlePhe 477
 DB 2223 GGTACTTCTATGCTACTCTCAATGTGACAGTAACTTGCCAAATTAAGGAGCACTTT 2282
 QY 478 ValLysAenAspGlyIleThrProLysProSerLeuLysValAlaLeuIleAlaGly 497
 DB 2283 GTAAATAATAGAGGGATACCTCTAAGCTTCCCTTTAAMAAGCTCTTAAATTCAGGT 2342
 QY 498 AlaAlaAspValGlyLeuGlyTrpProAenGlyAenGlnGlyTrpGlyArgValThrLeu 517
 DB 2343 GCTGGGAGATGTTGCACTTGGCTTCCAAATGATGTAACAAGAGGGGAAGATACGTTA 2402
 QY 518 AspLysSerLeuAenValAlaArgValAenGluSerSerAlaLeuSerThrSerGlnLys 537
 DB 2403 GATAAATCCCTAATATCGCATTTGGAATGAACAGAGCCCTTATCAACAAGCTCAAAA 2462
 QY 538 AlaThrThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTrpSer 557
 DB 2463 GCAACATATTCGTTACGCTCAAGCTGGTAAACCTTTAATAATATCACTGTTGGTCA 2522
 QY 558 AspAlaProAlaSerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIle 577
 DB 2523 GATGACCAAGTAGCAAGAGGATCACTTAATGTAATGATTAAGCTTAAGTAATC 2582
 QY 578 ThrAlaProAenGlyThrArgTrpValGlyAenAspPheSerAlaProPheAspAenAsn 597
 DB 2583 ACTGCAACAAATGGAATCAATATACGTCGGAATAATGACTTTACAGACACGATATGATACAA 2642
 QY 598 TrpAspGlyArgAenAenValGluAenValPheIleAsnSerProLysSerGlyThrTrp 617

Db 2643 TGGGATGCGAAGAAACAGCTGAGAAAATGCTTTATCATGCTCTCTCAAGCGGAGCTAT 2702
Qy 618 ThrileglValAlaIlnAlaIryAenValProValGlyProGlnAsnPheserLeuAlaIle 637
Db 2703 AAGGTGGAAGTGCAGGCTTACATGTAACGTAAGTCCGCAACCTTTCCTTACGAGATT 2762
Qy 638 ValAsn 639
Db 2763 GTRCAT 2768

RESULT 8
US-08-894-818B-2
; Sequence 2, Application US/08894818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikumoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-MAY-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03253
; FILING DATE: 07-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 323285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TAKAKURA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1977 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; US-08-894-818B-2

Alignment Scores:
Pred. No.: 9.76e-37 Length: 1977
Score: 477.00 Matches: 168
Percent Similarity: 41.84 Conservative: 90
Best Local Similarity: 27.24 Mismatches: 203
Query Match: 14.64 Indels: 156
Db: 3 Gaps: 25

US-10-784-870-4 (1-639) x US-08-894-818B-2 (1-1977)
Qy 62 ValGlySerGluAsnValIleLeuLeu---LysGlyLeuLeuIleLeuGlnThrVal 80

Db 103 GTTCAGCAGAAATACTACGACTGCTGACCCCGGAGCTCTTCAAGAAAGTCCAGAGATG 162
Qy 81 ProAlaAsnAsnIleLeuHilIleValGlnPheangIlyProIleLeuGlnIuThrIlys 100
Db 163 AACTGGAACCAAGAAAGTGCACCGCTCATTAATGTCGGG----- 201
Qy 101 GlnIlySleuGlnuThrThrcIyAlaIyAlaIleLeuApyrIleProApyrAlaIyIle 120
Db 202 -----AGCTAC--- 207
Qy 121 ValGluYrGluGlyAspValGlnserIyValArgserIleGlnHilIleValGlnserVal 140
Db 208 -----GAGACAGGAGCAGCGCGGTAAAGTACGTACAGCTCAGTACGGCCCGCAG 255
Qy 141 GlnProYrLeuProIySryIleIleAspProGlnLeuPheThr-----LysGlyAla 158
Db 256 GTCAAGTAC-----TCCTAACAAGATTAATCCCTGCTGCGCGTTAAATAAAGCCAGG 309
Qy 159 SerThrIleuValIyAlaIleuAlaIleuApyrThrIySgln-----AsnAsnIySglnVal 176
Db 310 GACCTTCTGCTGATGCGCGGCGCATGATACACAGGCTTACTTCGCTTAACAACAAGGCTTCG 369
Qy 177 GlnLeuArgGlyIleGluGlnIleAlaGlnIyValAlaIaserAsnAspValHilIyIle 196
Db 370 GGCATTAAGTTCATACAGAG----- 390
Qy 197 ThrAlaIySProGlnIyYrIyValIleAsnAspValAlaArgIyIle-----ValIyS 214
Db 391 -----GATTACAAAGGTTTCAGGTTGACGAGCCACTTCCTCCGACCATAGAGG 438
Qy 215 AlaAspValAlaGlnSerSerYrGlyLeuYrGlyGlnIyGlnIleValAlaValAla 234
Db 439 GCGCATACCGTGTGAGTACCTCTCGGCTTACACAGAGAGCGGTGCTGATGCTCCTC 498
Qy 235 AspThrGlyLeuAspThrGlyYrArgAsnAspSerSerMetHilIleValAlaPheArgGlyIyS 254
Db 499 GATTACGGGTATAGACGAGAC-----CACCCGAGTCTGAAGGCGAG 540
Qy 255 IleThrAlaLeuYr---AlaLeuGlyYrThrAsnAsnAlaAsnAspThrAsnGlyHilS 273
Db 541 GTCATAGGCTGGGTACGACGCGCTCAACGCGAGTTCAGCCCTTACATGACACGAGAC 600
Qy 274 GlyThrHilSValAlaGlySerValLeuGlyYrAsnGlyIleThrAsnIyS-----Gly 290
Db 601 GGAACCCAGTGTGCGGATGCTTGCAGAACCGGACGCTTAACCTCCACATGACATAGGC 660
Qy 291 MetAlaProGlnAlaAsnLeuValPheGlnSerIleMet-----AspSerSerGlyIyS 308
Db 661 GTCCGCCCGCGGAGAACTGTCGGCGCTCAAGGTTTCGGTCCGACGAGTTCGGGAGAC 720
Qy 309 LeuGlyIyLeuProSerAsnLeuGlnIleThrIlePheSerGlnAlaPheSerAlaGlyAla 328
Db 721 GTCCTCACCATCATCGCGGCTGTGACTGCGTCCGACAAACAAGACAAAGTACCGCATTA 780
Qy 329 ArgIle-----HilThrAsnSerIyGly 336
Db 781 AGGATCATCAACTTCTCCCTGCGCTCTCCAGAGCTCCAGAGAAACGACTCCCTCAGT 840
Qy 337 AlaAlaValaAsnGlyAlaIyYrThrThrAspSerArgAsnValaAspAspYrValaArgIyS 356
Db 841 CAGGCGCTCAACAACGCTGGAGCC----- 867
Qy 357 AsnAspMetThrIleLeuPheAlaAlaGlyYrAsnGlyYrProAangIyYrThrIleSer 376
Db 868 ---GGTATAGTATGTCGTGCGCGCGGCAACAGCGGCGGCAACACTTACACCTTCGCG 924
Qy 377 AlaProGlyThrAlaIyAsnAlaIleThrValGlyAlaThrGlnAsnLeuArgProSer 396
Db 925 TCACCCGCGCGCGGAGAGAGTCAATACCGTGGTGA----- 963
Qy 397 PheGlySerYrAlaAspAsnIleAsnHilValaIleGlnPheSerSerArgGlyProThr 416


```

Db      964 -----GTTGACAGCAAGCAACATCGCCAGCTTCTCCAGCAGGAGCGACC 1011
Qy      417 LysAspGlyValArgIleValProAspValMetAlaProGlyThrTyrIleLeuSerAlaArg 436
      1012 GCCGACGAGAGGCTCAAGCCGGAAGTGTGCGCCCGCTTGACATCAATAGTCCCGCGC
Qy      437 SerSerLeuAlaProAspSerSerPheTyrAlaAsnHisAspSerLysTyrAlaTyrMet 456
      1072 GCCAGC-----GGAACACGATGAGGACCCCGGATTAACGACTACACCAAGGCC 1122
Qy      457 GlyGlyThrSerMetAlaThrProIleValAlaGly---AsnValAlaGlnLeuArgGlu 475
      1123 TCTGGAACGACGATGCGCACCGCGAGTTTCGGCGCTTGAGCGGCTCATCTCCAGGCC 1182
Qy      476 HisPheValIleAsnArgGlyIleThrProLys-----ProSerLeuLeuValAla 493
      1183 CAC-----CCGAGCTGACCCCGGACGCGGACGAGGAGGAGACCGCC 1218
Qy      494 LeuIleAlaGlyAla-----AlaAspValGlyLeuGlyTyr 505
      1219 CTCATCGAGACCGCCGACATATGCGCCCAAGAGATAGCGGACATCGCTACGATGCGC 1278
Qy      506 ProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeu-----Asn 522
      1279 -----GGTAGGTGACGTCTACAGGCGCATCAAGCATCAAGTACGACGAC 1317
Qy      523 ValAlaTyrValLeuGlnSerSerAlaLeuSerThrSerGlnValAlaThrTyrPhe 542
      1318 TAGCGCCAGCTCACTTACCGGCTCGCTCGCGGACGAGGAGCGGACCGCACACCTTC 1377
Qy      543 ThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAlaSer 562
      1378 GACGTGACCGCGCCACCTTGTGACCGCACCTCTACTGAGAC----- 1422
Qy      563 ThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGly 582
      1423 -----ACGGGCTCGAGGACATCGACCTTACCTTACGACCGGAGG 1467
Qy      583 ThrArgTyrValGlyLeuAsnAspPheSerAlaProPheAspAsnThrAspGlyValAsn 602
      1468 AACGAG-----GTTGACTTACTCTTACACCGGCTACTAC 1500
Qy      603 AsnValGlnAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluValGln 622
      1501 GCGTTCGAGAGAGGCTGCTACTACACCGGACCGGAGGCTGAGCGGTACAGGTGTC 1560
Qy      623 AlaTyrAsnValProValGlyProGlnAsnAspPheSerLeuAlaIleValAsn 639
      1561 AGCTTACAG-----GGCGGCGGAGACTACAGGTGACGTCGTCAAC 1602
Db

```

```

RESULT 9
US-09-445-472-11
; Sequence 11, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMODO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

OTHER INFORMATION: Synthetic

US-09-445-472-11

Alignment Scores:	Length:
Pred. No.:	1977
Score:	477.00
Percent Similarity:	41.8%
Best Local Similarity:	27.2%
Query Match:	14.6%
DB:	3
Gaps:	25

US-10-784-870-4 (1-639) x US-09-445-472-11 (1-1977)

```

Qy      62 ValGlnSerGlnAsnValValLysLeuLeu---LysGlyLeuLeuLysLysLeuGluThrVal 80
      103 GTTCAGCAGAGAGACTACGAGCTGTCGACCCCGGAGCTGTTCAAGAAAGTCCAGAGATG 162
Qy      81 ProAlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlnGluThrLys 100
      163 AACTGAAACCGAGAGATGACACCTCATTAATGTTGCG----- 201
Qy      101 GlnLysLeuGlnThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIle 120
      202 -----ACGTAC--- 207
Qy      121 ValGluTyrGluGlyAspValGlnSerLysValArgSerIleGluHisValGlnSerVal 140
      208 -----GGAACAGGAGCAGGCGGTTAAGTACTGAGGCTCAGGCGCCAG 255
Qy      141 GluProTyrLeuProLysTyrLysIleAspProGlnLeuPheThr-----LysGlyAla 158
      256 GTCAAGTAC-----TCTTACAGATTAATCCTGCTGTCGCGGTTAAATTAAGCCAGG 309
Qy      159 SerThrLeuValLysAlaLeuAlaLeuAspThrLysGln-----AsnAsnLysGluVal 176
      310 GACCTTCTGCTGATCGCGGCGCATGATAGACACGGTTACTTCCGTAAACAGAGGTCGTCG 369
Qy      177 GlnLeuArgGlyIleGluGluIleAlaGlnTyrValAlaIleSerAsnAspValHisTyrIle 196
      370 GGCATTAAGTCTCTACAGGAG----- 390
Qy      197 ThrAlaLysProGluTyrLysValMetAsnAspValAlaArgGlyIle-----ValLys 214
      391 -----GATTACAGGTTTCAGGTTGACGACCCACTTCCTCCGATAGGG 438
Qy      215 AlaAspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnIleValAlaValAla 234
      439 GCGGATACCGTCTGAACTCCCTCGGCTACGAGGAGCGGTGTGGTGTGCGATCGTC 498
Qy      235 AspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGluAlaPheAspGlyLys 254
      499 GATACCGGTATTAACCGGAG-----CACCCGATCTGAAGGGCGAG 540
Qy      255 IleThrAlaLeuLys---AlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHis 273
      541 GTCATAGGCTGTGACACGCGCTCAACGCGAGGTCAACCCCTACATGACAGGAGAC 600
Qy      274 GlyThrHisValAlaGlySerValLeuGlyAsnGlyAlaThrAsnLys-----Gly 290
      601 GGAACCAAGTTCGGGTATCGTTGCGGAAACCGGAGGTTAATCCCACTACATAGGC 660
Qy      291 MetAlaProGlnAlaAsnLeuValPheGlnSerIleMet-----AspSerSerGlyGly 308
      661 GTCCGCCCGCGCGAGAGCTCGTCCGCTCAAGGTTCTCGTCCGACGAGTTGCGGAGGC 720
Qy      309 LeuGlyLysLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAla 328
      721 GTTCCACATCATCGCGGCTGTGACTGGGTGTCGAGAACAGAGCAAGTACGGGATA 780
Qy      329 ArgIle-----HisThrAsnSerTyrGly 336
      781 AGGGTATTAACCTTCCCTCGGCTCTCCAGAGCTCGAGGAGGACGACTCTCCAGT 840

```



```

QY 337 AAlaValAsnGlyValThrThrAspSerArgValAspArgValArgLys 356
DB 841 CAGCGCCGACCAAGCCCTGGACGCC----- 867
QY 357 AsnAspMetThrIleLeuPheAlaAlaGlyAsnGluArgProAsnGlyLysThrIleSer 376
DB 868 ---GGTATGATGTCGCGTCGCGCGGCAAGCGGCGCAACCTACACCGTCGCGC 924
QY 377 AlaProGlyThrAlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSer 396
DB 925 TCACCGCGCGCGCGCAAGCTCATACCGTCGTCGCA----- 963
QY 397 PheGlySerThrAlaAspAsnIleAsnIleValAlaGlyPheSerSerArgGlyProThr 416
DB 964 -----GTTGACAGCAACAGCAACATCGCCAGCTTCTCCAGCAGGGAGCCGAC 1011
QY 417 LysAspGlyArgGlyLeuArgProAspValMetAlaProGlyThrThrIleLeuSerAlaArg 436
DB 1012 GCGGACGGAAAGCTCAAGCGCGAAGTCGTGCGCCCGCGGCTTGACATCAGCCCGCGC 1071
QY 437 SerSerLeuAlaProAspSerPheThrAlaAsnIleAspSerLysThrAlaThrMet 456
DB 1072 GCGACG-----GAAACAGCATGGGCAACCCGATAAAGACTACTACCAAGGCC 1122
QY 457 GlyGlyThrSerMetAlaThrProIleValAlaGly---AsnValAlaGlyLeuArgGly 475
DB 1123 TCTGGAAACGATGGCCACCCGCGACGTTTGGGGGCTGGCGCTCCTCCAGGCC 1182
QY 476 HisPheValLysAsnArgGlyLysLeuProLys-----ProSerLeuLeuLysAlaAla 493
DB 1183 CAC-----CCGAGCTGGACCCCGCAACAGGTGAAGACCGCGC 1218
QY 494 LeuIleAlaGlyAla-----AlaAspValGlyLeuGlyThr 505
DB 1219 CTCATCGAAGCCCGCACTAGTCCGCCCAAGAGATGCGGACCTCGCTCAAGCTGCG 1278
QY 506 ProAsnGlyAsnGlnGlyTrpGlyArgValThrLeuAspLysSerLeu-----Asn 522
DB 1279 -----GCTAGCGTGAACGTCTACAGGCGCATCAAGTACAGCAGAC 1317
QY 523 ValAlaThrValAsnGluSerSerAlaLeuSerThrSerGlnLysAlaThrThrPhe 542
DB 1318 TAGCGCAAGCTCACTTACCGGCTCCGCGCGCAAGAGAGGCGCCACCTTC 1377
QY 543 ThrAlaThrAlaGlyLysProLeuLysIleSerLeuValIleProSerAspAlaProLaser 562
DB 1378 GACGTCAGGCGCGCCACTTCTGTGACCGCCACTCTACTGGGAC----- 1422
QY 563 ThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGly 582
DB 1423 -----ACGGGCTCGAAGCATCGACCTTACTCTACGACCCGAC 1467
QY 583 ThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnIleTrpAspGlyArgAsn 602
DB 1468 AACGAG-----GTTGACTACTCTCTACACCCGCTACTAC 1500
QY 603 AsnValGluAsnValPheIleAsnSerProGlnSerGlyThrThrIleGluValGln 622
DB 1501 GCGTTCGAGAGAGTGGCTACCAACCGGACCGGAACTGGACGCTCAAGGTCGTC 1560
QY 623 AlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
DB 1561 AGCTAFAAG-----GGCGCGGCAACTACAGGTGACGTGTGACG 1602

```

RESULT 10

```

US-10-090-624-11
; Sequence 11, Application US/10090624
; Patent No. 6783970
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMODO, Tomoko
; APPLICANT: ASADA, Kiyoko

```

```

; APPLICANT: KATO, Ikunobu
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 11
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; US-10-090-624-11

Alignment Scores:
Pred. No.: 9,766-37 Length: 1977
Score: 477.00 Matches: 168
Percent Similarity: 41.8% Conservative: 90
Best Local Similarity: 27.2% Mismatches: 203
Query Match: 14.6% Indels: 156
DB: 3 Gaps: 25

US-10-784-870-4 (1-639) x US-10-090-624-11 (1-1977)
QY 62 ValGluSerGluAsnValLysLeuLeu---LysGlyLeuLeuLysLeuGluThrVal 80
DB 103 GTTCAGCAGAGAACTACGAGCTGTCGACCCCGGACCTGTTCAAGAAAGTCCAGAGATG 162
QY 81 ProAlaAsnLysLysLeuIleValGlnPheAsnGlyProIleLeuGluGluThrLys 100
DB 163 AACTGGAAACGAGAGGTGACACCGCTCAATGTTCCGG----- 201
QY 101 GlnLysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIle 120
DB 202 -----AGCTAC--- 207
QY 121 ValGluTyrGluGlyAspValGlnSerLysValArgSerIleGluHisValGluSerVal 140
DB 208 -----GGAGACAGGAGACAGCGCGGTTAAAGTACGAGGCTATGCGGCCGAC 255
QY 141 GluProTyrLeuProLysTyrLysIleAspProGlnLeuPheThr-----LysGlyAla 158
DB 256 GTCAAGTAC-----TCTTCAAGATATATCCCTGTCGCGGTTAAATTAAGGCCAG 309
QY 159 SerThrLeuValLysAlaLeuAlaLeuAspThrLysGln-----AsnAsnLysGluVal 176
DB 310 GACCTTCTGCTAGTCGCGGCGATGATGACAGCGGTTACTTCGTTACACAAAGGTCGCG 369
QY 177 GlnLeuArgGlyIleGluGluIleAlaGlnTyrValAlaSerAsnAspValIleTyrIle 196
DB 370 GGCATTAAGTTCATACAGAG----- 390
QY 197 ThrAlaLysProGluTyrLysValMetAsnAspValAlaArgGlyIle-----ValLys 214
DB 391 -----GATTCAAGGTTCAAGTTACACACGCTTCCTCCGACAGATGGG 438
QY 215 AlaAspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAla 234
DB 439 GCCATACCGTCTGGAATCCTCGGCTACGACGAAAGCGGTGTGGTGGTCCATCGTC 498
QY 235 AspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGluAlaPheArgGlyLys 254
DB 499 GATACGGGTATACAGCGGAGC-----CACCCGATCTGAAGGGCAG 540
QY 255 IleThrAlaLeuTyr---AlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHis 273
DB 541 GTCATAGGCTGTGACGAGCGCGTCAAGCGCAGGTGACCCCTTCGATGACCAAGGACAC 600

```

```

QY 274 G1YThrh1vAla1ag1ySerVal1eug1yAaNg1yAlaThraNlyS-----G1Y 290
DB 601 GGAACCCACCGTGGGATGCTGGCCGGAACCGGATTACTCCACATACATAGGC 660
QY 291 MetAlaProGlnAlaAenLeuValPheGlnSer11Met-----AapSerSerg1y 308
DB 661 GTGCCCCCGGCGGAAGCTGCGGCGTCACAGGTTCTCGTGCAGCGGTTCCGGAAGC 720
QY 309 LeuG1yG1LeuProSerAenLeuGlnThrLeuPheSerg1AlaPheSerAlG1yAla 328
DB 721 GTCTCCACCATCATCGGGGTGTGACTGGGTCCTCCAGAACAGACAGATACGGGATA 780
QY 329 Arg1Le-----HisThraSerTrpG1y 336
DB 781 AGGCTCATCAACCTCTCCCTGCGCTCTCCAGAGCTCCAGCGAATCCGATCTCTCACT 840
QY 337 AlaAlaValAaNg1yAla1yTrhThraPserArgAaValAaPaeP1yValArg1yS 356
DB 841 CAGGCCCTCAACAACGCTCGGACGCC----- 867
QY 357 AaNaPaeThrl1eLeuPheAla1ag1yAaNg1yAaArgProAaNg1yG1yThrl1eSer 376
DB 868 ---GGTAAGTAAAGTCTGCGTGGCGCGCAACGCGGCAACCTTACACCGTGGC 924
QY 377 AlaProG1yThra1a1yAaNaAla1eThraValG1yAlaThrg1uAaenLeuArgProSer 396
DB 925 TCACCCGCGCGCGACCAAGTCAACCTCGGTCA----- 963
QY 397 PheG1ySerTyAlaAaPaeN11eAaN11eValA1aGlnPheSerSerg1yProThr 416
DB 964 -----GTTGACAGCAACAGACATCGCGCTTCTCAGACGAGGACCGACC 1011
QY 417 LysAaNg1yArg1le1ySProAaPValMetAlaProG1yThry1leleuSerAlaArg 436
DB 1012 GCGAGCAAGGCTCAAGCCGGAAGTGTGCGCGCGCTTGAATCATAGACCGCGCG 1071
QY 437 SerS1rLeuAlaProAaPserSergPheTrpAlaAaN1aPaeSerg1yAla1yMet 456
DB 1072 GCCAGC-----GGAACACAGATGGGACCCCGCAATAACATCTACACCAAGGC 1122
QY 457 G1yG1yThrSerMetAlaThrPro1leValA1ag1y---AaValA1aGlnLeuArg1u 475
DB 1123 TCTGGAACAGCATAGCCACCGCGACGTTGCGGCGTGGCGGCTCATCTCCAGGCC 1182
QY 476 HisPheVal1yAaAaNg1y1leThrPro1yS-----ProSerLeuLeu1yAla1a 493
DB 1183 CAC-----CCGAGCTGACCCCGGACAGGAGGAGAGCCGCC 1218
QY 494 Leu1leAlaG1yAla-----AlaAaPValG1yLeuG1yTy 505
DB 1219 CTATTCAGACCGCCGACATAGTCGCCCCCAAGAGATAGCGGACATCGCTACGGTGGC 1278
QY 506 ProAaNg1yAaNg1yTrpG1yArgVal1ThrLeuAaP1ySerLeu-----Aa 522
DB 1279 -----GGTAGGCTGACGTCACAGGCAATCAAGCAATCAAGTACAGAC 1317
QY 523 ValAlaTyValAaen1ySerSerAlaLeuSerThrSerg1yAla1yThry1Phe 542
DB 1318 TAGCGAAGCTCACTTCAACGCGCTCGCGCAGACAGGAGGAGCGCACCCACCTTC 1377
QY 543 ThrAlaThra1aG1ySProLeu1y1leSerLeuVal1yPserAaPAlaProAlaSer 562
DB 1378 GACGTCAAGCGCGCACCTTCGTGACCGCACCTTACTGAGAC----- 1422
QY 563 ThrThra1aSerVal1ThrLeuValAaNaPae1yAaPLeuVal1leThra1aProAaNg1y 582
DB 1423 -----ACGGGCTCGAGGAGCATGCACTTCACTCACTACGACCCCAAGCGG 1467
QY 583 ThrArgTyVal1yAaenAaPheSerAlaProPheAaPaeAaen1yPaeP1yArgAa 602
DB 1468 AACGAG-----GTTGACTAATCTCTACACCGGCTTACTAC 1500
QY 603 AaNaValG1uAaenValPhe1leAaenSerProGlnSerg1yThry1Thrl1eG1uValGln 622

```

```

DB 1501 GGCTTCAGAAAGGTGCGCTACTACACCCGCGGAACTGAGCGTCAAGTGTGC 1560
QY 623 AlaTyAaNaValProValG1yProGlnAaPheSerLeuAla1eValAa 639
DB 1561 AGCTACAAAG-----GGCGCGCGAGAACTACACAGTCAAGCTGCTGACG 1602

RESULT 11
US-09-841-553-2
; Sequence 2, Application US/09841553
; Patent No. 6849441
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; MORISHITA, Mio
; YAMAMOTO, Katsuhiko
; MITTA, Masanori
; ASADA, Kiyozo
; TSUNASAWA, Suenmu
; KATO, Ikunobu
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; City: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,553
; FILING DATE: 24-Apr-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/894,818
; FILING DATE: <Unknown>
; APPLICATION NUMBER: JP 323285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TAKAKURA=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1977 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-841-553-2

Alignment Scores:
Prod. No.: 9,766-37 Length: 1977
Score: 477.00 Matches: 168
Percent Similarity: 41.8% Conservative: 90
Best Local Similarity: 27.2% Mismatches: 203
Query Match: 14.6% Indels: 156
DB: 3 Gaps: 25

US-10-784-870-4 (1-639) x US-09-841-553-2 (1-1977)
QY 62 valG1uSerG1uAaenVal1ySLeuLeu---LyG1yLeuLeu1ySLeuG1uThraVal 80
DB 103 GTTCACAGAAAGAACTACGAGCTGCTGACCGCGGAGCTTTCAGAAAGTCCAGAGAG 162
QY 81 ProAlaAaNaen1ySLeu1leVal1eValGlnPheAaNg1yPro1leleuG1uThraVal 100

```

DB 163 AACTGAAACCAAGAGTGAACCGCTCATATATGTTGGG----- 201
QY 101 GlnLysLeuGlnThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrMet 120
DB 202 -----ACTAC----- 207
QY 121 ValGluTyrGluGlyAspValGlnSerLysValArgSerIleGluHisValGluSerVal 140
DB 208 -----GGAAACAGAGGACAGGCGGTTAAGGTACGTACAGGCTCATGGCGCCGAG 255
QY 141 GluProTyrLeuProLysIleAspProGlnLeuPheThr-----LysGlyAla 158
DB 256 GTCAGATAC-----TCTTACAGATATATCTCTGCTGCTGCGGTAAATTAAGGCCAG 309
QY 159 SerThrLeuValLysAlaLeuAlaLeuAspThrLysGln-----AsnAsnLysGluVal 176
DB 310 GACCTTCTGCTGATCGCGGCGCATGATAGACACGGGTACTTCCGTAAACACAGGCTCTCG 369
QY 177 GlnLeuArgGlyIleGluGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIle 196
DB 370 GGCATTAAGTCTATACGAG----- 390
QY 197 ThrAlaLysProGluTyrLysValMetAsnAspValAlaArgGlyIle-----ValLys 214
DB 391 -----GATTACAAAGTTCAAGTTGACGACGCCACTTCCGCTCCACGATAGGG 438
QY 215 AlaAspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnIleValAlaValAla 234
DB 439 GCGATACCGCTGGAATCTCCCTCGGCTACGACGAAAGCGGTGTGGTGGTCATCGTC 498
QY 235 AspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGluAlaPheArgGlyLys 254
DB 499 GATACGGGTATACACCGCAAC-----CACCCGATCTGAAGGGCAAG 540
QY 255 IleThrAlaLeuTyr-----AlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHis 273
DB 541 GTCATAGGCTGGTACGACCGCTGCAACGCGAGGTCCACCCCTACATACACGAGGACAC 600
QY 274 GlyThrHisValAlaGlySerValLeuGlyAsnGlyAlaThrAsnLys-----Gly 290
DB 601 GGAAACCAAGTTCGGGGTATCGTTGCCGGAACCGGACGGTTATCCCAAGTACATAGGC 660
QY 291 MetAlaProGlnAlaAsnLeuValPheGlnSerIleMet-----AspSerSerGlyGly 308
DB 661 GTCGCGCCCGCGGCAAGCTCTCGCGGCTCAAGGTTCTCGTGCACGATTCGGGAAAC 720
QY 309 LeuGlyLysLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAla 328
DB 721 GTCCTCAACATCATCGCGGGTGTGACTGGTCTCCAGAAACAGACAAAGTACGGGATA 780
QY 329 ArgIle-----HisThrAsnSerTyrGly 336
DB 781 AGGGTATCAACCTCTCCCTCGGCTCTCCAGAGCTCCAGACGAAACGATCCCTCACT 840
QY 337 AlaAlaValAsnGlyAlaTyrThrThrAspSerArgAsnValAspAspTyrValaLysLys 356
DB 841 CAGGCGCGTCAACACCGCTGGAGCGCC----- 867
QY 357 AsnAspMetThrIleLeuPheAlaAlaGlyAsnGluArgProAsnGlyLysThrIleSer 376
DB 868 ---GGATATGATGCTGCTGCTCGCGCGGCAACAGCGGCGCAACCTACACCTGCGGC 924
QY 377 AlaProGlyThrAlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSer 396
DB 925 TCACCGCGCGCGGCAAGGTCAATACCTCGGGTCA----- 963
QY 397 PheGlySerTyrAlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThr 416
DB 964 -----GTTGACAGCAACGACATCGCCAGGCTTCTCCAGCGGAGACGAGAC 1011
QY 417 LysAspGlyArgIleLysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArg 436

DB 1012 GCGACAGAAAGCTCAAGCCGGAAGTCTGCGCCCGCGGTGACATCATACCCCGCGC 1071
QY 437 SerSerLeuAlaProAspSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMet 456
DB 1072 GCCAGC-----GGAACAGCATGGGACCCCGATTAAGACTACTACACCAAGGCC 1122
QY 457 GlyGlyThrSerMetAlaThrProIleValAlaGly-----AsnValAlaGlnLeuArgGlu 475
DB 1123 TCTGAACCAAGATGGCCACCCCGCATGTTTGCGGCGTTGGCGCGCTCATCTCCAGGCC 1182
QY 476 HisPheValLysAsnArgGlyIleThrProLys-----ProSerLeuLeuLysAla 493
DB 1183 CAC-----CCGACTGACCCCGGACAGAGTAAAGCCGCC 1218
QY 494 IleuIleAlaGlyAla-----AlaAspValGlyLeuGlyTyr 505
DB 1219 CTCATCGAGACCGCGACATAGTCCGCCCAAGAGATGCGGACATCGCTCATCGGTGCG 1278
QY 506 ProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeu-----Asn 522
DB 1279 -----GTAGGGTGAACGTTACMAAGCCATCAAGTACAGACGAC 1317
QY 523 ValAlaTyrValAsnGluSerSerAlaLeuSerThrSerGlnLysAlaThrTyrThrPhe 542
DB 1318 TAGCCCAAGCTACCTTCAACCGGCTCCGTCGCCGCAAGGAAGGCCACCACTTC 1377
QY 543 ThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrPheAspAlaProLaser 562
DB 1378 GACGTACAGCGGCGCACCTTCGTACCGCCACCTCTACTGGGAC----- 1422
QY 563 ThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGly 582
DB 1423 -----ACGGGCTCGAGCGACATCACTCTACTTACGACCCCAACGGG 1467
QY 583 ThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAspTyrAspGlyArgAsn 602
DB 1468 AACGAG-----GTTGACTACTCTTACACCGGCTCATAC 1500
QY 603 AsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluValGln 622
DB 1501 GCGTTCGAGAGGTGCGCTACTACCAACCGGACCGCGGAACTGGAAGGTCAAGGTCTGC 1560
QY 623 AlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
DB 1561 ACCTACAG-----GGCGGCGAATCAACGATCGACGTCTGAC 1602
RESULT 12
US-08-894-818B-34
Sequence 34, Application US/08894818B
Patent No. 6261822
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

QY 577 ILeThrAlaProkenglyThrArgTyValGlyAsnAspPheSerAlaProPheAspAsn 596
DB 1450 CTCCTAGATCCCATGGAACGAG-----GTTGACTTAC 1482
QY 597 AsnTrpAspGlyArgAsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThr 616
DB 1483 TCTTACACCGCCTACATGAGATTGGAAGGTGGTTATTACACCACTGATGGAACA 1542
QY 617 TyrThrIleGluValGlnAlaIleValArgAsnValProValGlyProGlnAsnPheSerLeuAla 636
DB 1543 TGGACAAATTAAAGGTGTGATGCTACAGC-----GGAAGTCAAACTATCAATGATGAT 1593
QY 637 IleValAsn 639
DB 1594 GTGGTAAGT 1602
RESULT 13
US-09-445-472-15
Sequence 15, Application US/09445472
Patent No. 6358726
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hi karu
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Kunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/09/445,472
CURRENT FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 15
LENGTH: 1962
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-445-472-15 .
Alignment Scores:
Pred. No.: 2,116-33 Length: 1962
Score: 443.00 Matches: 163
Percent Similarity: 40.3% Conservative: 88
Best Local Similarity: 26.2% Mismatches: 204
Query Match: 13.6% Indels: 168
DB: 3 Gaps: 25
US-10-784-870-4 (1-639) x US-09-445-472-15 (1-1962)
QY 62 ValGluSerGluAsnValIleuLeuLeuGlyLeuLeuValLeuLeuGluThrValPro 81
DB 103 GTTGAAGAAGACTATAGTGTCTGCTAACGCGAGCTGTTCAAAAATTCAAAATTTGAT 162
QY 82 AlaAsnAsnValLeuHisIleValGlnPheAsnGlyProIleLeuGlnGluThrIleGln 101
DB 163 CCTAACGAGAAATCAGCAGCTA-----ATTGTAATTTGAAAAACCATAGGGGAA 210
QY 102 Lys-----LeuGluThrThrGlyAlaIleValIleLeuAspTyrIlePro 115
DB 211 AAAAGAAATTGCGTAGAGACTTCTTAGATTAAATGGGGCAAAAGTT----- 255
QY 116 AspTyrAlaTyrIleValGluTyrGluGluAspValGlnSerIleValArgSerIleGlu 135
DB 256 -----AGGTATGCTG----- 264
QY 136 HisValGluSerValGluProTyrIleuProTyrIleAspProGlnLeuPheThr 155
DB 265 -----TACCATAATTATACCGCAGATA----- 285

QY 156 LysGlyAlaSerThrLeuValIleuValAlaLeuAspThrIleGlnAsnAsnValGlu 175
DB 286 ---GCTCCGATCTTAAGTTTAGAGCTTACATGATCTCAGTTTAAACAGGGGTAA 342
QY 176 ValGlnLeuArgGlyIleGluGluIleAlaGlnTyrValAlaSerAsnAspValIleTyr 195
DB 343 GCTAAGCTTTCAAGGTTAGGTTTATC-----CAGAAACATACAA 384
QY 196 ILeThrAlaLysProGluTyrLysValIleAsnAspValAlaArgGlyIleValIleVal 215
DB 385 GTTACAGTTTCAGCAAAATTAGAAAGACTGATGATGCTGCAGCTCAAGTTATGCAACT 444
QY 216 AspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnGluIleValAlaValAlaAsp 235
DB 445 TACGTTTGAACTTGGGATATGAT-----GATTCTGAATCAATATGAATTAATGAC 498
QY 236 ThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGluAlaPheArgGlyLysIle 255
DB 499 ACTGGAATTGAC-----GCTTCATCCAGATCTCAAGAAAGTA 540
QY 256 ThrAlaLeuTyrAlaLeuGlyArgThrAsnAsnAlaAsn-----Asp 269
DB 541 -----ATTGGGCGGTAGATTGTCATATGTTAGGAGATTATCATATGAT 585
QY 270 ThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGlyAsnGlyValaThrAsn 288
DB 586 GACCATGACATGGAATCTCATATGATCTCATATGACAGCTGGATCTGAGACAGCAATGAT 645
QY 289 -----LysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMet----- 303
DB 646 GCGAAGTCAAGGAAATGGCTCCAGAGCTAAGCTGGGGGAATTAAAGTTCTAGTGCC 705
QY 304 AspSerSerGlyIleGluGlyLeuProSerAsnLeuGlnThrIleuPheSerGlnAla 323
DB 706 GATGTTCTCGAAGCATATCTACTTAATTAAAGGAGTGAAGTGGCCGTTGATTAACA 765
QY 324 PheSerAlaGlyAlaArgIleHisThrAsnSerTyrGlyAla----- 337
DB 766 GATTAAGTACGAAATTAAGCTCTTAATCTTCTTGGTTCAAGCAGATCTCAGATAGT 825
QY 338 -----AlaValAsnGlyValaTyrThrAspSerArgAsnValAsp 351
DB 826 ACTGACGCTTAAGTCAGAGCTGTTAATGACAGTGGATGCT----- 867
QY 352 AspTyrValaGlyLysAsnAspMetThrIleLeuPheAlaAlaGlyAsnGluArgProAsn 371
DB 868 -----GATTAAGTTGTTGTTGGTGGCCGCGGAAACAGTGAACCTTAAC 909
QY 372 GlyIleThrIleSerAlaProGlyThrAlaLysAsnAlaIleThrValGlyValaThrGlu 391
DB 910 AAGTATACATCGGTTCTCCAGCAGCTCCAGCAAAAGTTATTTACGTTGAGCC----- 963
QY 392 AsnLeuArgProSerPheGlySerTyrAlaAsnAsnIleAsnHisValAlaGlnPheSer 411
DB 964 -----CTTGAACAAGTATGATGTTATTAACAAGCTTCA 996
QY 412 SerArgGlyProThrLysAspGlyArgIleLeuProAspValIleMetAlaProGlyThrTyr 431
DB 997 AGCAGAGGCGCAATCGCAGAGCGAGCTTAAGCTGAGTTGTTGCTCCAGAAACCTGG 1056
QY 432 IleLeuSerAlaArgSerSerLeuAlaProAspSerSerPheThrAlaAsnHisAspSer 451
DB 1057 ATAATGCTGCCAGAGCAAGT-----GGAATACAGATGGGTCAAACATTAATATAC 1107
QY 452 LysTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIleValAlaGlyAsnValAla 471
DB 1108 TATTACACAGCAGCTCCGAGACATCAATGCAACTCTCAAGTACGCTGATATGCAAGC 1167
QY 472 GlnLeuArgGluHisPheValIleAsnArgGlyIleThrProLys-----ProSerLeu 489
DB 1168 CTCTTGCTCCAA-----GCACACCCGAGCTGAGCTCCAGACAA 1206
QY 490 LeuLysAlaIleAlaLeuIleAlaGlyAla-----AlaAspVal 501

```

Db      1207  GTAAACAGCCCTCATGAAACGCTGATATCGTAAAGCAGATGAATAATGCCGATATA 1266
Qy      502  G1yeucllyrProenglyangllyrpolylarvalThrleuaplyserleu 521
Db      1267  GCCTAGCGTGA-----GTAAGGTTAAATGCATACAGGCTATA 1305
Qy      522  AsnValAlaTyrValaanglyserSeraleuSerThrserglnlyala----- 538
Db      1306  AAC-----TTCGTAATCTATGCAGAGCTAGTTCCTGATATGTTGCCAACAAAGGC 1359
Qy      539  -----ThyTyrThrphrThrAlaThrAlaGlyProleuyls1eserleuValTyr 556
Db      1360  AGCCAACTCAACGATTCGTTATTAAGCGAGCTTCGTTCCGACACTTATCTACGG 1419
Qy      557  SerAspAlaProalaserThrThrAlaSerValThrleuValaAsnAapleuAapleuVal 576
Db      1420  GACATGCCAAT-----ACCGACTTGATCTTTAC 1449
Qy      577  IleThrAlaProaanglyThrArgTyrValGlyAsnAapPheSerAlaProPheAspAsn 596
Db      1450  CTCACATCCCAATGGAACAG-----GTTGACTAC 1482
Qy      597  AsnTTPaapGlyArgaAsnValGluAsnValPheIleAsnSerProGlnAsnSerGlyThr 616
Db      1483  TCTTACACCGCTACTATGATTCGAAAGCTTCGTTATACACCACTCATGATGACA 1542
Qy      617  TyrThrIleGluValGlnAlaTyrAsnValProValGlyProGlnAapPheSerleuAla 636
Db      1543  TGACCAATTAAAGCTTGTAGCTACAG-----GGAAGTCAACTATCACTAGAT 1593
Qy      637  IleValaAsn 639
Db      1594  GTGCTAAGT 1602

RESULT 14
US-10-090-624-15
; Sequence 15, Application US/10090624
; Patent No. 6783970
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1962
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-090-624-15

Alignment Scores:
Pred. No.: 2,11e-33 Length: 1962
Score: 443.00 Matches: 163
Percent Similarity: 40.3% Conservative: 88
Best Local Similarity: 26.2% Mismatches: 204
Query Match: 13.6% Indels: 168
DB: 3 Gaps: 25

US-10-784-870-4 (1-639) x US-10-090-624-15 (1-1962)

```

```

Qy      62  ValGluSerGluAsnValValylleuLeuLeuGlyleuLeuylslyleuGluThrValPro 81
Db      103  GTTGAAGAAACCTATGCTCTGCTAACCCAGAGACTGTTCCAGAAAATTCAAAATTTGAT 162
Qy      82  AlaAsnAsnlySleuHisIleValGlnPheAsnGlyProIleLeuGlnGluThrlyleuGln 101
Db      163  CCTAACGAGGAATACACACAGTA-----ATTGATTTGAAAACCAATAGGAA 210
Qy      102  Lys-----LeuGluThrGlyValAlaValylleuAapTyrIlePro 115
Db      211  AAAGAAATTCACATTAAGAGTTCTTGAAGTAAAGGCTCAAAAGTT----- 255
Qy      116  AspTyrAlaTyrIleValGluTyrGluGlyAspValGlnSerlySvalArgSerIleGlu 135
Db      256  -----AGTATGG----- 264
Qy      136  HisValGluSerValGluProTyrLeuProlyrlySylleAapProGlnLeuPheThr 155
Db      265  -----TACCATTTATACCCGCAAT----- 285
Qy      156  LysGlyAlaSerThrleuVallyValAlaLeuAlaAsnThrlySglnAaAsnlySgln 175
Db      286  ---GTCGCCGATCTTAAGGTTAGAGACTTACTGATCTCAGGTTTAAAGGGGCTAAA 342
Qy      176  ValGlnleuArgGlyIleGluGluIleAlaGlnTyrValAlaSerAsnAapValHisTyr 195
Db      343  GCTAAGCTTTCAGGTGTAGGTTATTC-----CAGAAAGCTACAAA 384
Qy      196  IleThrAlaAsnProGluTyrlySvalMetAsnAapValAlaArgGlyIleValAla 215
Db      385  GTTACAGTTTCACAGCAATTAAGAAGACTGATGATCTCAGCTCAAGTTATAGCAACT 444
Qy      216  AspValAlaGlnSerSerTyrGlyleuTyrGlyGlnGlyGlnIleValAlaValAlaAap 235
Db      445  TACGTTTGAACTTGGAATATGAT-----GTTCTGGAATCACAATAGGAATATATGAC 498
Qy      236  ThrGlyleuAsnThrGlyArgaAsnAapSerSerMetHisGluAlaPheArgGlylySle 255
Db      499  ACTGAATTTGAC-----GCTTCTCATCCAGATCTCCAAAGAAAGTA 540
Qy      256  ThrAlaLeuTyrAlaLeuGlyArgThrAsnAsnAlaAsn-----Aap 269
Db      541  -----ATTGGTGGTAAAGATTTTGTCAATGTAGCACTTCCATACGAT 585
Qy      270  ThrAsnGlyAlaGlyThrHisValAlaGlySerValleuGlyAsnGlyAlaThrAsn--- 288
Db      586  GACCATGACATGAACTCATGTAGCTTCAATAGCAGCTGTTCTGAGCAGCAAGATAT 645
Qy      289  -----LysGlyMetAlaProGlnAlaAsnleuValPheGlnSerIleMet----- 303
Db      646  GGCAGATACAAAGGAATGCTCCAGAGACTTAAGCTGCGGGAATTAAGGTTCTAGGTGCC 705
Qy      304  AapSerSerGlyValylleuGlyGlyleuProSerAsnleuGlnThrleuPheSerGlnAla 323
Db      706  GATGTTCTGGAAGCATATCTACTATTAATTAAGCGAGTGAAGTGGCCGCTGATTAACAA 765
Qy      324  PheSerAlaGlyAlaArgIleHisThrAsnSerTyrGlyAla----- 337
Db      766  GATTAAGTACGAATTAAGCATTTATCTTTCTCTGTTCAAGCAGAGCTCAGATGCT 825
Qy      338  -----AlaValaanglyAlaTyrThrThrAspSerArgAsnValAap 351
Db      826  ACTGACGCTTAAGTCAGGCTGTTAATGCAAGCGCTGGATGCT----- 867
Qy      352  AapTyrValAlaArglyAsnAapPheThrIleleuPheAlaAlaGlyAsnGluArgProAsn 371
Db      868  -----GATTAAGTTGTTGGTTGGCTGGAACAGTGAACCTTAAC 909
Qy      372  GlyGlyThrIleSerAlaProGlyThrAlaValAsnAlaIleThrValGlyAlaThrGlu 391
Db      910  AAGTATACAGATCGTTCTCCAGAGCTGCAAGCAAGTTATTAAGTTGAGCC----- 963
Qy      392  AsnleuArgProSerThrleuGlySerTyrAlaAsnAapHisValAlaGlnPheSer 411

```



```

Db 964 -----GTTGACCAAGTATGATTAATACCAAGCTTCTCA 996
Qy 412 SERTAGGLYPROTHIRLYSAPGLYARGILEYSPROAPVALMELALAPROGLYTHR 431
    |||||
Db 997 ACCAGAGGGCCAACTGACAGACGAGCTTAAGCTGAGGTGGTCTCAGAGAACTGG 1056
    |||||
Qy 432 ILEUSERALARGSERSERLEUALAPROAPSERPHERPALAAMHISAPSER 451
    |||||
Db 1057 ATRATGCTGCCAGGCAAGT-----GGAACTAGAGATGGGTCAACCAATATATGAC 1107
    |||||
Qy 452 LYSYTRALATYRMEGLYLYTHRSEMERLATHPRILLEVALAAGLYSVALA 471
    |||||
Db 1108 TATTACACAGACGCTCGGACATCAATGGCAACTCTCACTGATCGTATTTGACGC 1167
    |||||
Qy 472 GINLEUARGGLIHISPEVALYSAAMARGGLYILETHPROLYS-----PROSERLEU 489
    |||||
Db 1168 CTCTTCTCCAA-----GCACACCCGAGCTGACCTCCAGCAAA 1206
    |||||
Qy 490 LEULYALALALEULILEAGLYALA-----ALAPVAL 501
    |||||
Db 1207 GTRAAAACAGCCCTCATAGAACTGCTGATATGTAAGCAGATGAATAAGCCGATATA 1266
    |||||
Qy 502 GLYLEUGLYTYRPROAENGLYANGINGLYTRPGLYARGVALTHIRLEUAPLYSERLEU 521
    |||||
Db 1267 GCCTACGCTGCA-----GTTAGGATTATGCAATACAAAGCTATA 1305
    |||||
Qy 522 AANVALALATYRVALANGUSERSERALALEUSERTHSERGLNLYSALA----- 538
    |||||
Db 1306 AAC-----TACATATACATGCAAGCTAGCTTCACTGATATGTCACCAAGGC 1359
    |||||
Qy 539 -----THTYRTHRPHETHRALATHRALAGLYLYSPROLEULYSILESERLEUVALITTP 556
    |||||
Db 1360 AGCCAACTCAGCAGCTCGTTATTAAGCGAGCTTCGTTACATGCGCAATTAATACG 1419
    |||||
Qy 557 SERAPALAPROALASERTHRTHRALASERVALTHIRLEUALAMHAPLEUAPLEUVAL 576
    |||||
Db 1420 GACAAATGCCAAT-----AGCACTTGAATCTTTAC 1449
    |||||
Qy 577 ILETHRALAPROAENGLYTHRARGTYRVALGLYASAPSPHESERLAPROPEAPAPAN 596
    |||||
Db 1450 CTCTACGATCCCAATGGAACAG-----GTTGACTAC 1482
    |||||
Qy 597 AANTTPAPGLYARGAMENVALGLUAMNVALPHEILEAMSERPROGLINSERGLYTHR 616
    |||||
Db 1483 TCTTACACCCCTACTATGATTGCAAAAGTGGTATTATTAACAACCACTGATGAAACA 1542
    |||||
Qy 617 TYRTHIRLEGLIUALGLALATYRASNVALPROVALGLYPROGLINAMPHESERLEUALA 636
    |||||
Db 1543 TGGCAATTAAAGTTTGAAGCTACAC-----GGAAGTGCAAACTATCAAGTATGAT 1593
    |||||
Qy 637 ILEVALAAN 639
    |||||
Db 1594 GTGTTAGT 1602
    |||||

```

```

STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,553
FILING DATE: 24-Apr-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/894,818
FILING DATE: <unknown>
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1962 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-841-553-34
Alignment Scores:
Pred. No.: 2,116-33 Length: 1962
Score: 443.00 Matches: 163
Percent Similarity: 40.3% Conservative: 88
Best Local Similarity: 26.2% Mismatches: 204
Query Match: 13.6% Indels: 168
Gaps: 25
US-10-784-870-4 (1-639) x US-09-841-553-34 (1-1962)
Qy 62 VALIGUSERGLUASNVALLYSEULEUYSGLYLEUULYLYLEUGLUTHRYVALPRO 81
    |||||
Db 103 GTTGAGAAACTATAGCTGTCTTAAGCCAGACCTGTCAAGAAAATTCAAAAATTGAAT 162
    |||||
Qy 82 ALAAMENLYLEUHIILEVALGLINPHEANGLYPROILEUGLUGLUTHRYSGIN 101
    |||||
Db 163 CCTAACGAGAAATCAGACAGTA-----ATTGTTTGGAAACCATAGGGGA 210
    |||||
Qy 102 LYS-----LEUGLUTHRTHRGIALALYALILEUAMPTYRILEPRO 115
    |||||
Db 211 AAAGAAATTGCACTAAGACTTCTTGATTATGGGTGCAAAAGTT----- 255
    |||||
Qy 116 APTYRALATYRILEVALIURGLUGLYASPVALGINSELYSVALARGSERILEGLU 135
    |||||
Db 256 -----AGGTATGTC----- 264
    |||||
Qy 136 HSAVALIGUSERVALGLUPROFYRLEUPROLYTYRILEASPPOGLINLEUPHERTHR 155
    |||||
Db 265 -----TACCATATTATACCGCAATA----- 285
    |||||
Qy 156 LYSGLYALASERTHRLEUVALYSALEUALALEUASPTRHLYSGINAMHAPLYSGIN 175
    |||||
Db 286 ---GCTGCCGATCTTAAGTTAGAGCTTACATGATCTCAGGTTAAACAGGGGTAAA 342
    |||||
Qy 176 VALGINLEUARGGLYILEGLUGLUILEAGLINTYRVALALASERAMHAPVALHISLYR 195
    |||||
Db 343 GCTAACCTTTCAGGTGTTAGTTATC-----CAGGAAGACTGCAAA 384
    |||||
Qy 196 ILETHRALALYSPROGLIUTHRYVALMETAMENAPVALALARGGLYILEVALYSALE 215
    |||||

```


Db 385 GTTACAGTTTCAGCAGAAATTAGAAAGACTGAGTAGCTGACAGCTCAAGTTATGCAACT 444
 Qy 216 AepValAlaGlnSerSerThrGlyLeuThrGlyGlnGlnIleValAlaValAlaAsp 235
 Db 445 TACGTTTGAACTTGGGATATGAT-----GGTCTGGAATACACATAGGAATATATAC 498
 Qy 236 ThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGlnAlaPheArgGlyLeuIle 255
 Db 499 ACTGGAATTGAC-----GCTTCTCATCCAGATCTCCAGAGAAAGTA 540
 Qy 256 ThrAlaLeuThrAlaLeuGlyArgThrAsnAlaAsn-----Asp 269
 Db 541 -----ATTGGGTGGGTAGATTGTCATGTAAGGAGTTATCCATACAT 585
 Qy 270 ThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGlyAsnGlyAlaThrAsn--- 288
 Db 586 GACCATGACATGAACTCATGATGATCTCAATAGCAAGCTGTGATCTGGAGCAGCAATGAT 645
 Qy 289 -----LysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMet----- 303
 Db 646 GCGAAGTACAAAGGAATGGCTCCAGAGCTAAGCTGGCGGAATTAAGTTCTAGTCC 705
 Qy 304 AspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAla 323
 Db 706 GATGTTCTGGAAACATATCTACTATATTAAGGAGTGAAGTGGCCGTTGATTAACA 765
 Qy 324 PheSerAlaGlyAlaArgIleHisThrAsnSerThrGlyAla----- 337
 Db 766 GATTAAGTACGGAATTAAGTCATTAATCTTCTCTGTTCAACGACAGCTCAGATGCT 825
 Qy 338 -----AlaValAsnGlyAlaThrThrAspSerArgAsnValAsp 351
 Db 826 ACTGACCGCTTAAGCTGCTTTAATGCAAGCTGGAGTCT----- 867
 Qy 352 AspThrValArgIlyAsnAspMetThrIleuPheAlaAlaGlyAsnGlyArgProAsn 371
 Db 868 -----GGATGATGTTGTTGGTGGCTGGGAAACAGTGAACCTTAC 909
 Qy 372 GlyGlyThrIleSerAlaProGlyThrAlaLeuAsnAlaIleThrValGlyAlaThrGlu 391
 Db 910 AAGTATCAATCGGTTCTCCAGCAGCTGCAGCAAGTTATTACAGTTGGAGCC----- 963
 Qy 392 AsnLeuArgProSerPheGlySerThrAlaAspAsnIleAsnIleValAlaGlnPheSer 411
 Db 964 -----GTTGACAAAGTATGATGTTTATTAACAAGCTTCTCA 996
 Qy 412 SerArgIlyProThrIlyAspGlyArgIleIleuPheAspValMetAlaProGlyThrTyr 431
 Db 997 AGCAGAGGCGCAACTGACAGCGGCAAGCTTAAGCTGATGTTGCTCCAGAAACTGG 1056
 Qy 432 IleuSerAlaArgSerSerLeuAlaProAspSerPheThrAlaAsnHisAspSer 451
 Db 1057 ATTAATGCTGCCAGAGCAAGT-----GGAACCTAGCATGGGTCAACCAATTAATGAC 1107
 Qy 452 LysThrAlaThrMetGlyIleThrSerMetAlaThrProIleValAlaGlyAsnValAla 471
 Db 1108 TATTACACAGAGCTCTGGAGCATGCAATGCAACTCTCACTGAGCTGATTTGACAGCC 1167
 Qy 472 GlnLeuArgGlnHisPheValIlyAsnArgGlyIleThrProLys-----ProSerLeu 489
 Db 1168 CTCTGTGCTCAA-----GCACACCGGAGCTGAGCTCCAGACAA 1206
 Qy 490 LeuIlyPheAlaLeuIleAlaGlyAla-----AlaAspVal 501
 Db 1207 GTTAAACACGCTCATAGAACTGCTGATATGCTAAAGCCAGATGAATAGCCGATATA 1266
 Qy 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyThrGlyArgValThrLeuAspLysSerLeu 521
 Db 1267 GCTTACGCTGCA-----GTTAGGTTAATGCAATCAAGGCTATA 1305
 Qy 522 AsnValAlaThrValAsnGlnSerSerAlaLeuSerThrSerGlnIlyVala----- 538
 Db 1306 AAC-----TACGATTAATGCAAAAGTATGTTCTGATATGTTGCCAACAAGGC 1359

Qy 539 -----ThrThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTrp 556
 Db 1360 AGCCAACTCACAGTTCTGTTATTAAGGAGCTTCGTTGTAACCTGCACATTAATACG 1419
 Qy 557 SerAspAlaProAlaSerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuVal 576
 Db 1420 GACATATCCAT-----ACGACCTTGATCTTTAC 1449
 Qy 577 IleThrAlaProAsnGlyThrArgThrValGlyAsnAspPheSerAlaProPheAspAsn 596
 Db 1450 CTTAACGATCCCAATGGAACCAAG-----GTTGACTAC 1482
 Qy 597 AsnThrAspGlyArgAsnAsnValGlnAsnValPheIleAsnSerProGlnSerGlyThr 616
 Db 1483 TCTTACACCGCTACTATGATTCGAAAGGTTGGTTATTACAAACCAACTGATGAAACA 1542
 Qy 617 TyrThrIleGlyValGlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAla 636
 Db 1543 TGGCAATTTAAGGTTGTAAGCTACAGC-----GAAAGTGCAAACTATCAAGTAT 1593
 Qy 637 IleValAsn 639
 Db 1594 GTGTTAGT 1602

Search completed: April 8, 2006, 03:53:40
 Job time : 585 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: April 8, 2006, 03:55:21 ; Search time 1318 Seconds
(without alignments)
4009.205 Million cell updates/sec

Title: US-10-784-870-4

Perfect score: 3267
Sequence: 1 MRKKVFLSLTSAAILSTV.....EVQAVNPVGPQNSLAIVN 639

Scoring table:

BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-Q=/abseq/ABSSMB.spool/US10784870/runat_07042006_143427_27724/app_query.fasta_1
-DB=Published Applications NA Main -GEMT=fastap -SUPFX=g2n.trnblm
-MIMATCH=0.1 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1
-MATRIX=bls0um62 -TRANS=human40.cdt -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext
-HEP_SIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs804
-USRR=US10784870 @CCN 1.1 1026 @runat_07042006_143427_27724 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3267	100.0	1920	3	US-09-920-954-3
2	3267	100.0	1920	7	US-10-784-870-3
3	3072	94.0	1923	3	US-09-920-954-7
4	3072	94.0	1923	7	US-10-784-870-7
5	3066	93.8	1923	3	US-09-920-954-5
6	3066	93.8	1923	7	US-10-456-479-3
7	3066	93.8	1923	7	US-10-784-870-5

8	3066	93.8	1923	8	US-10-820-712A-2	Sequence 2, Appl1
9	3066	93.8	1923	9	US-10-820-714A-2	Sequence 2, Appl1
10	2237	68.5	1302	8	US-10-820-712A-11	Sequence 11, Appl1
11	2237	68.5	1302	9	US-10-820-714A-12	Sequence 12, Appl1
12	2188	67.0	1302	8	US-10-820-712A-13	Sequence 13, Appl1
13	2188	67.0	1302	9	US-10-820-714A-14	Sequence 14, Appl1
14	2183	66.8	1305	6	US-10-385-662-1	Sequence 1, Appl1
15	2122	65.0	1302	8	US-10-820-712A-21	Sequence 21, Appl1
16	2122	65.0	1302	9	US-10-820-714A-22	Sequence 22, Appl1
17	2002.5	61.3	1299	8	US-10-820-712A-19	Sequence 19, Appl1
18	2002.5	61.3	1299	9	US-10-820-714A-20	Sequence 20, Appl1
19	1998.5	61.2	1299	8	US-10-820-712A-15	Sequence 15, Appl1
20	1998.5	61.2	1299	9	US-10-820-714A-16	Sequence 16, Appl1
21	1991.5	61.0	1299	8	US-10-820-712A-17	Sequence 17, Appl1
22	1991.5	61.0	1299	9	US-10-820-714A-18	Sequence 18, Appl1
23	477	14.6	1977	3	US-09-841-553-2	Sequence 2, Appl1
24	477	14.6	1977	5	US-10-090-624-11	Sequence 11, Appl1
25	477	14.6	1977	8	US-10-800-684-2	Sequence 2, Appl1
26	477	14.6	1977	9	US-10-888-588-11	Sequence 11, Appl1
27	443	13.6	1962	3	US-09-841-553-14	Sequence 34, Appl1
28	443	13.6	1962	5	US-10-090-624-15	Sequence 15, Appl1
29	443	13.6	1962	8	US-10-800-684-14	Sequence 34, Appl1
30	443	13.6	1962	9	US-10-888-588-15	Sequence 15, Appl1
31	432	13.2	1977	3	US-09-841-553-6	Sequence 6, Appl1
32	432	13.2	1977	8	US-10-800-684-6	Sequence 6, Appl1
33	417.5	12.8	1236	5	US-10-090-624-2	Sequence 2, Appl1
34	417.5	12.8	1236	9	US-10-888-588-2	Sequence 2, Appl1
35	417.5	12.8	1566	3	US-09-841-553-4	Sequence 4, Appl1
36	417.5	12.8	1566	8	US-10-800-684-4	Sequence 4, Appl1
37	359	11.0	3624	6	US-10-156-761-5701	Sequence 5701, Ap
38	359	11.0	9025608	6	US-10-156-761-1	Sequence 1, Appl1
39	357.5	10.9	3417	6	US-10-156-761-3306	Sequence 3306, Ap
40	357.5	10.9	9025608	6	US-10-156-761-1	Sequence 1, Appl1
41	344.5	10.5	135638	9	US-10-314-657-1	Sequence 1, Appl1
42	344.5	10.5	135638	9	US-10-473-193-1	Sequence 1, Appl1
43	342.5	10.5	1739	3	US-09-974-300-1934	Sequence 1934, Ap
44	326	10.0	4765	3	US-09-841-553-7	Sequence 7, Appl1
45	326	10.0	4765	5	US-10-090-624-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-09-920-954-3
Sequence 3, Application US/09920954
Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKI, KATSUSHI
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-OPCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 1920
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURES:
NAME/KEY: CDS

/ LOCATION: (1) .. (1920)
US-09-920-954-3

Alignment Scores:

Pred. No.:	1,226-287	Length:	1920
Score:	3267.00	Matches:	639
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-784-870-4 (1-639) x US-09-920-954-3 (1-1920)

```

QY 1 MetArgbysbysbysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrVal 20
DB 1 ATGAGAAAGAAAGAGAGGCTTTTATCTGTTTATCAGCTGCGAGATTCGTGACGCT 60
QY 21 AleuLeuAenAnPseSerAlaGlyAspAlaArgThrPheApsLeuApsPheLySgLyLe 40
DB 61 GCATTAAACAATCCCTCGGCTGGTGATGCAAGGACTTTGATCTGATTTTAAAGGAAAT 120
QY 41 GluThrThrThrAspValSerGlyPheSerLyGlnArgGlnThrGlyAlaAlaIlePhe 60
DB 121 CAACAACAACAACGATGTCAGTGGTTCTCCAAACAGACAAACAAGTGGCTGCAATTT 180
QY 61 LeuValGluSerGluAenValLysLeuLeuLySgLyLeuLeuLySbysLeuGluThrVal 80
DB 181 CTGGTGAGCTGTAATAATGTAATGTAATCTTTAAAGATTCCTTAAAGAACTTGAACGTA 240
QY 81 ProAlaAenAnLysLeuHsIleValGlnPheAenGlyProIleLeuGluGluThrLys 100
DB 241 CCGGCAAAATATTAATCCCATATTTGTCATTCATTCATGCGCCCATTTTAAAGAAACAAA 300
QY 101 GluLysLeuGluThrThrGlyAlaLysIleLeuApsTrtIleProApsTrtAlaTrtIle 120
DB 301 CAAGAAGTAAAGCAACGAGCAAGAAATTCGACATCAATCCCTGATTAATGCAATTAAT 360
QY 121 ValGluTrtGluGlyAspValGlnSerLyValArgSerIleGluHsIleValGluSerVal 140
DB 361 GTGAGATAGAGGGGAGATGTTCACTGCAAAAGTCCTCCATTAACACGTGGAAATCAAGT 420
QY 141 GluProTrtLeuProLySbysTrtLysIleApsProGlnLeuPheThrLySgLyAlaSerThr 160
DB 421 GAGCCCTACTTGGCGAAATCAAAATAGATCCCACTTTTCAAAAAGGCGATGACGAG 480
QY 161 LeuValLysAlaLeuAlaLeuApsThrLysGlnApsAenLysGluValGlnLeuArgGly 180
DB 481 CTGGTAAAGCGTTGGCGCTGTATACAGACAGAACATTAAGAAATGCAATTAAGAGGC 540
QY 181 IleGluGluIleAlaGlnTrtValAlaSerApsApsValHsIleTrtIleThrAlaLysPro 200
DB 541 ATCGAAGGAATCGCTCAATGACGTAGCAAGCAATGACGTCAATTAATTAACGCAAGCCT 600
QY 201 GluTrtLysValMetApsApsValAlaArgGlyTrtValLysAlaApsValAlaGlnSer 220
DB 601 GAATATAAGGTGATGATGATGATGCGCAAGGATATGCAAAAGCGATGTGCAACGAGC 660
QY 221 SerTrtGlyLeuTrtGlyGlnGlyGlnIleValAlaValAlaApsThrLysLeuApsThr 240
DB 661 AGCTACGCTTTGATATGAGCAAGGCGCAGATGTGCGAGTTGCGATCTGAAATGGAATACA 720
QY 241 GlyArgApsApsSerSerMetHsIleGluAlaPheArgGlyLysIleThrAlaLeuTrtAla 260
DB 721 GGAAGAAACGACGATTGATGATGATGAGCTTCCGCGATTAATAACACACATTAATGCA 780
QY 261 LeuGlyArgThrApsAenAlaApsApsThrApsGlyHsIleGlyThrHsIleValAlaGlySer 280
DB 781 CTGGGTCCGACCAATTAATGCAATGATGACAAACGATGATGATGATGATGATGATGATG 840
QY 281 ValLeuGlyApsGlyAlaThrApsApsGlyMetAlaProGlnAlaApsLeuValPheGln 300
DB 841 GATTTAGAAATGGCGCAACGAATTAAGAAATGCGACCTCAAGCGAATCTGATTTTCA 900

```

```

QY 301 SerIleMetApsSerSerGlyGlyLeuGlyGlyLeuProSerApsLeuGlnThrLeuPhe 320
DB 901 TCCATCATGATATAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIleHsIleThrApsSerTrtGlyAlaAlaValAps 340
DB 961 AGCGAAGCATTCAGTGCAGATGCGCAAGATTCATCAAAATCTCTGGGGGCGACCGGTGAAT 1020
QY 341 GlyAlaTrtThrApsSerApsAenValApsApsTrtValArgLysApsApsMetThr 360
DB 1021 GGGGCTTACACACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 361 IleLeuPheAlaAlaGlyApsGlyLysArgProApsGlyGlyTrtIleSerAlaProGlyThr 380
DB 1081 ATCTTTTGGCGATGAGAAATGAGCGGAAACGCGGATCAATGACATGACATGATGATGATG 1140
QY 381 AlaLysApsAlaIleThrValGlyAlaThrGluApsLeuArgProSerPheGlySerTrt 400
DB 1141 GCTAAACACGCAATTAACGTCGCGCAACCGCAACCGCAACCGCAACCGCAACCGCAAC 1200
QY 401 AlaApsAenIleApsHsIleValAlaGlnPheSerSerArgGlyProThrLysApsGlyArg 420
DB 1201 GCAGATTAATATTAACACGTCGCAAGTCTCTCCGTCGCGCAACCAAAAGATGGCGCA 1260
QY 421 IleLysProApsValMetAlaProGlyThrTrtIleLeuSerAlaArgSerSerLeuAla 440
DB 1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 441 ProApsSerSerPheTrtAlaApsHsIleApsSerLySbysTrtAlaTrtLysTrtSer 460
DB 1321 CCGGATTCCTCTCTGCGCGCAATCAACACGCAAAATATGCTTAATGATGATGATGATGAT 1380
QY 461 MetAlaTrtProIleValAlaGlyApsValAlaGlnLeuApsGlyHsIlePheValLysAps 480
DB 1381 ATGGCAACACCGATTTGTCGCGGAAATGTCACAGTCGCGGATTTTGTGAATAAT 1440
QY 481 ArgGlyIleThrProLysProSerLeuLysValAlaLeuIleAlaGlyAlaAlaAps 500
DB 1441 AGAGAAATCACTCTTAAGCTTCTCTTAATGAAGAGCTTTGATGATGATGATGATGATGAT 1500
QY 501 ValGlyLeuGlyTrtProApsGlyApsGlnGlyTrtGlyArgValThrLeuApsLysSer 520
DB 1501 GTTGATTTGGGTTATCCGAAACGAAACCAAGATGCGGCGGATGATGATGATGATGATGAT 1560
QY 521 LeuApsValAlaTrtValAlaApsGlySerSerAlaLeuSerThrSerGlnLysAlaThrTrt 540
DB 1561 TTGAACGTTGCTTAATGTAAGAAATCCAGTCCCTATCACTAGCCAAAGCGACATAT 1620
QY 541 ThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTrtPseApsAlaPro 560
DB 1621 AACTTTACTGCAACGCGCGGCAAGCATTTGAATTTGATTTGATTTGATTTGATTTGATTT 1680
QY 561 AlaSerThrThrAlaSerValThrLeuValApsApsLeuApsLeuValIleThrAlaPro 580
DB 1681 GCAAGCACTACTGCTTGTGAACCTGGGTCAATGATTTGATTTGATTTGATTTGATTTGAT 1740
QY 581 ApsGlyThrArgTrtValGlyApsApsPheSerAlaProPheApsApsApsTrtApsGly 600
DB 1741 AACGGAACAAATATGTCGGAATGATCTTCAACACATTTGCAATTAACCTGGAGATGGC 1800
QY 601 ArgApsApsValGluApsValPheIleApsSerProGlnSerGlyTrtTrtThrIleGlu 620
DB 1801 CGCAATTAACGTAAGAAATGATTTTAATTTATTTGATTTGATTTGATTTGATTTGATTTGAT 1860
QY 621 ValGlnAlaTrtApsValProValGlyProGlnApsPheSerLeuAlaIleValAps 639
DB 1861 GTGCAAGCATTAATGTCGCGGTGACCAACAAATCTCTGTTGGCAATGATGATGATGATG 1917

```

RESULT 2

US-10-784-870-3

; Sequence 3, Application US/10784870

; Publication No. US20040142837A1

; GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKAI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKURA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784, 870
PRIORITY FILING DATE: 2004-02-24
PRIORITY APPLICATION NUMBER: US/09/509, 814A
PRIORITY FILING DATE: 2000-04-06
PRIORITY APPLICATION NUMBER: PCT/JP98/04528
PRIORITY FILING DATE: 1998-10-07
PRIORITY APPLICATION NUMBER: JP 9-274570
PRIORITY FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 1920
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
US-10-784-870-3

Alignment Scores:

Pred. No.:	1,22e-287	Length:	1920
Score:	3267.00	Matches:	639
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-784-870-4 (1-639) x US-10-784-870-3 (1-1920)

QY 1 MetArgLysLeuValPheLeuSerValLeuSerAlaAlaAlaLeuSerThrVal 20
DB 1 ATAGAGAAAGAAAGAGGTGTTTATCTGTTTATCAGCTCAGCGATTCTGCGACTGTT 60
QY 21 AlaLeuAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyLe 40
DB 61 GCATTAAACATCTCTCGCTGATGATGCAAGACCTTTGATCTGGAATTTAAAGGAATT 120
QY 41 GlnThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaPhe 60
DB 121 CAACCAACACCATGATGATGCTGTTCTCCAAACAGGACAAACAGGTGGCTGCATTT 180
QY 61 LeuValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLysLeuGluThrVal 80
DB 181 CTGGTGAAGCTGAAAATGTGAACCTTCTTAAGAGATTGCTAAAGAACTTGAACAGTA 240
QY 81 ProAlaAsnAsnLysLeuHisGlnGlnPheAsnGlyProIleLeuGluGluThrLys 100
DB 241 CCGGCAAAATATAATCACTCCATATGTCTCAATTCATATGCCCATTTTGAAGAAACAAA 300
QY 101 GlnLysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaLysTyrIle 120
DB 301 CAGACGTAAGACCACTGAGCAAGCAAGATTCTGCACTACATCCCTGATATAGCATATATT 360
QY 121 ValGluTyrGluGlyAspValGlnSerLysValArgSerIleGluHisValGlnSerVal 140
DB 361 GTGAGATATGAGGGGAGTTCAGTCAAAAGTCGCTCCATTGAACAGTGAATCAGTG 420
QY 141 GlnProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThr 160
DB 421 GAACCACTACTTGCAGAAATCAAAATAGATCCCAAGCTTTTCAAAAAGCGCATCAGC 480
QY 161 LeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgGly 180

DB 481 CTGGTGAAGAGCTTGCGCTTGATACGAAGCAGAACATAAAGATGCAATTAAGAGC 540
QY 181 IleGluGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaLysPro 200
DB 541 ATCAGAGAAATCCCTCAGTACGTAGCAAGCAAGACCTCCATATATATTACGGCAACCT 600
QY 201 GlnTyrLysValIleCysAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSer 220
DB 601 GAATATAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 221 SerTyrGlyLeuTyrGlyGlnGlnIleValAlaValAlaAspThrGlyLeuAspThr 240
DB 661 AGCTACGCTTGTATGAGCAAGGCAAGTATGTCAGTTCGCAATCTGATGATGATGATGATGAT 720
QY 241 GlyArgAsnAspSerSerMetHisGluAlaPheArgGlyLysIleThrAlaLeuTyrAla 260
DB 721 GGAAGAAACGACGTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 261 LeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySer 280
DB 781 CTGGTGAAGAGATATATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 281 ValLeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGln 300
DB 841 GTATTGGAATGCGCGCAACGAATTAAGAAATGCGACCTCAAGCAATCTGTTTTTCAA 900
QY 301 SerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLysPhe 320
DB 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyrGlyAlaAlaValAsn 340
DB 961 AGCCAAAGCATTCAGTCAAGGTGCAAGAAATTCATCAAACTCTCGGGGGGCAAGCGTGAAT 1020
QY 341 GlyAlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThr 360
DB 1021 GGGGCTTACACGACATTCCTCAAAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 361 IleLeuPheAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThr 380
DB 1081 ATCTTTTCGCGCTGGAATGAAAGCGGAGCAAGCGGTGATCAGTGCACCTGATGAC 1140
QY 381 AlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyr 400
DB 1141 GCTAAAGAGCCCAATACAGTGGCGCAACCGAAACCTGGTCCAGCTTCGATCTCTAT 1200
QY 401 AlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArg 420
DB 1201 GCAGATATATTAACAAGTTCACAGTTCTTCCCGTGGCCGCAAAAAGATGGCGCA 1260
QY 421 IleLysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAla 440
DB 1261 ATCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 441 ProAspSerSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMetGlyGlyThrSer 460
DB 1321 CCGGATTCCTCTTCTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 461 MetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPheValLysAsn 480
DB 1381 ATGGCAACACCATTTTTCGGGAAATGTTGCAACAGCTCCGTGACATTTTGGAAAAAT 1440
QY 481 ArgGlyIleThrProLysProSerLeuLeuLysAlaAlaLeuIleAlaGlyAlaAlaAsp 500
DB 1441 AGAGAAATCACTCTCAAGCTTCCTCATTTGAAGCAGCTTGAATGATGATGATGATGATGATGAT 1500
QY 501 ValGlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSer 520
DB 1501 GTTGAATGGGTTATCCGAAACGAAACCAAGATGGGGCGCAAGTACCTCGGATTAATATCG 1560
QY 521 LeuAsnValAlaTyrValAsnGluSerSerAlaLeuSerThrSerGlnLysAlaThrTyr 540
DB 1561 TTGAACGTTGCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620

QY 541 ThrPheThrAlaThrAlaGlyLeuProLeuValIleSerLeuValTTPSerAlaPro 560
DB 1621 ACCCTTAACGCAACGGGGGCAAGCCATTAATAATCCCTGGATATGGTGGATGCCCTT 1680
QY 561 AlaSerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaPro 580
DB 1681 GCAGACCTACCTGCTTCTGTACCCCTGGTCAATGATTGGCATTAACGACCA 1740
QY 581 AenGlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTPaArgly 600
DB 1741 AACGGAACAGATATATGCGGAATGACTTCTCAGCACCATTTGACATATACCTGGATGCG 1800
QY 601 ArgAsnAspValGlyAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGln 620
DB 1801 CGGAATTAACGTAAATAATGATTATTAATTCGCCCAAGTGAACATATACCATTTGAG 1860
QY 621 ValGlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuValIleValAsn 639
DB 1861 GTGCAAGCATATATATGCGCGTTGGACCAACAAATCTTCTGCTGGCATTTGTGAC 1917

RESULT 3
US-09-920-954-7
Sequence 7, Application US/09920954
Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKIWA, MIKO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-OPCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-09-920-954-7

Alignment Scores:
Pred. No.: 7,58e-270 Length: 1923
Score: 3072.00 Matches: 596
Percent Similarity: 97.0% Conservative: 23
Best Local Similarity: 93.4% Mismatches: 19
Query Match: 94.0% Indels: 0
DB: 3 Gaps: 0

US-10-784-870-4 (1-639) x US-09-920-954-7 (1-1923)

QY 2 ArgGlyValGlyValPheLeuSerValLeuSerAlaAlaIleLeuSerThrValAla 21
DB 7 AAGAAGAAAAAGGTGTTTATCTGTTTATCATCGTCGACGATTTTGTGCACTGTGGC 66
QY 22 LeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLeuGlyIleGln 41
DB 67 TTAAGTAATCCATCGCAGGTGTGTCAGAAATTTTGATCGAATTTCAAGAAATTCAG 126

QY 42 ThrThrThrAspValSerGlyPheSerIleAsnArgGlnThrGlyAlaAlaPheLeu 61
DB 127 ACAACAACTGAATGCTAAAGGTTTCTCCAGACGGGGCACAATGCTGCTCTTTCG 186
QY 62 ValGlySerGlyAsnValIleLeuLeuLeuGlyLeuLeuValLeuGlnThrValPro 81
DB 187 GTGGAATCGAAAATGTGAACCTCCCAAAAGGTTTGCAAGAAAGACCTTGAACAGTCCCG 246
QY 82 AlaAsnAsnLeuLeuIleValGlnPheAsnGlyProIleLeuGlnGlnThrLyGln 101
DB 247 GCAAAATTAATTAATCCATATTTATCCAAATTCATATGACCAATTTTGAAGAAACAAACAG 306
QY 102 LysLeuGlnThrThrGlyValAlaValIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
DB 307 CAGCTGAAAAAACAGGGGCAAAAGATTCGACTACATACCTGATTAATGCTTAACATGTC 366
QY 122 GlnTyrGlnGlyAspValGlnSerIleValAsnSerIleGlnIleValGlnSerValGln 141
DB 367 GAGTATGAGGGCGGATTTAAGTCACGACACAGCACCATTTAGACGTGAAATCCGTGGAG 426
QY 142 ProTyrLeuProLysTyrIleAspProGlnLeuPheThrLysGlyValaSerThrLeu 161
DB 427 CCTTATTTGCCGATATACAGATATAGATCCCGCTTTTCAAAAGGGGCATCAGAGCTT 486
QY 162 ValIleValIleValIleAspThrIleGlnAsnAsnValGlnValGlnLeuArgIle 181
DB 487 GTAAAGACGTGGCTTGATACAAAGCAAAATTAAGAGTCAATTAAGAGCGATC 546
QY 182 GlnGlnIleAlaGlnTyrValAlaSerAsnAspValIleTyrIleThrAlaLysProGln 201
DB 547 GAAACAATCGCAATTCGCAATTAAGCAATATGATGCTATATATACGGAAGCTGAG 606
QY 202 TyrIleValMetAsnAspValAlaArgGlyIleValIleValAlaAspValAlaGlnSer 221
DB 607 TATAAGTGATGAATGAATGTTGCGGTGAATTCGAAAGCGAGTGTGCTCAGAGCAGC 666
QY 222 TyrGlyLeuTyrGlyGlnGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
DB 667 TACGGGTGTATGACAAAGACAGATCGTAGCGGTGCGAATACAGGGCTTGATACAGGT 726
QY 242 ArgAsnAspSerSerMetIleGlnAlaPheArgIleLysIleThrAlaLeuTyrAlaLeu 261
DB 727 CGCAATGACAGTTGATGATGATGAGCTTCCGCGGAAATTAATGATTAATGATGATG 786
QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyIleGlyThrIleValAlaGlySerVal 281
DB 787 GACCGGACGAATTAAGCCAAATGATGATGATGATGATGATGATGATGATGATGATGATG 846
QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
DB 847 TTAGAAACGGCTCCATATTAAGAAATGGCGCTCAGGCGAATCTTGTCTTCCAAATCT 906
QY 302 IleMetAspSerSerGlyIleGlyLeuGlyIleLeuProSerAsnLeuGlnThrLeuPheSer 321
DB 907 ATCATGATAGTCGGGCGGAGCTTGAGAGCATACCTTGGAATTCGAAACCTTATTCAGC 966
QY 322 GlnAlaPheSerAlaGlyAlaArgIleIleIleThrAsnSerTTPGlyAlaAlaValAsnGly 341
DB 967 CAAGCATCAAGTGTGCGGCGAGAAATTCATCAAACTCTGGGAGCGAGGTGAATGGG 1026
QY 342 AlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThrIle 361
DB 1027 GCTTACACACAGATTCAGAAATGTGATATCATATGCGCCAAATATGATATGACATC 1086
QY 362 LeuPheAlaAlaGlyAsnGlyArgProAsnGlyIleThrIleSerAlaProGlyThrAla 381
DB 1087 CTTTTCGCTGCGGGAATGAAGACCGAAACGGCGGAACATCAAGTGCACGAGGACAGCT 1146
QY 382 LysAsnAlaIleThrValGlyAlaThrGlyAsnLeuAspProSerPheGlySerTyrAla 401
DB 1147 AAAAATGCAATTAACGTGCGAGCTTACGGAACCTCCGCGCAACCTTTGGGTCTTATGCG 1206
QY 402 AspAsnIleAsnIleValAlaGlnPheSerSerArgIleProThrLysAspGlyArgIle 421

Db 1207 GACAAATTCACACATGTCGACAGTTCCTTCACTGACGACCAAGATGACGATC 1266
Qy 422 LysProAspValMetAlaProGlyThrTrpIleLeuSerAlaArgSerSerLeuAlaPro 441
Db 1267 AACCGAATGTCATGCGACCGGAAACGTTCACTATCAAGAAATCTTCTTTCGACCG 1326
Qy 442 AspSerSerPheTrpAlaAsnHisAspSerIleTrpAlaTrpMetGlyGlyThrSerMet 461
Db 1327 GATTCCTCTTCTTGGGGAACATGACAGTAAATATGATATCAATGCGGACATCCATG 1386
Qy 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlnHisPheValIleAsnArg 481
Db 1387 GCTACACGATCGTTGTCGAAACGTGACAGCTTCGTGACATTTTGTGAAAAACGA 1446
Qy 482 GlyIleThrProIleProSerLeuLeuValAlaAlaLeuIleAlaGlyValAlaAspVal 501
Db 1447 GGCATTCACACCAAGACCTTCTCTATTAACCGGACATGATTCGCGGCTGACATC 1506
Qy 502 GlyLeuGlyThrProAsnGlyAsnGlnGlyTrpGlyArgValThrLeuAspIleSerLeu 521
Db 1507 GGCCTTGGCTACCCGAAACGATACCAAGATGGGACAGTGAATGATTAATCCCTG 1566
Qy 522 AsnValAlaTrpValAsnGlySerSerAlaLeuSerThrSerGlnIleAlaThrTrpThr 541
Db 1567 AACCTTCCTATGTAAGACAGTCCAGTTCTCTATCCACAGCAAAAGACGATCTCG 1626
Qy 542 PheThrAlaThrAlaGlyIleProLeuIleIleSerLeuValTrpSerAspAlaProAla 561
Db 1627 TTACTGCTACTGCGGCAAGCTTTGAAATATCCCTGATGATGTCGATGACCCCTCG 1686
Qy 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCAACACTCTTCCTCGTAAGCTTGTCAATGATTTGACCTTGTCAATACCGCTCCAAAT 1746
Qy 582 GlyThrArgTrpValGlyAsnAspPheSerAlaLeuSerThrSerGlnIleAlaThrArg 601
Db 1747 GGCACACAGTATGAGAAATGACTTATCTTCGCAATACATGATTAATCGGAGATGCGCG 1806
Qy 602 AsnAsnValGlnLeuAsnValPheIleAsnSerProGlnSerGlyThrTrpIleGlnVal 621
Db 1807 AATACAGTAAAGAAATGATTAATTAATGACCAACAAAGCGGACGTTATACATTTGAAGTA 1866
Qy 622 GlnAlaIleTrpAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGCTTAATACGATCCGTTGACCAAGAACTTCTGTTGGCAATTTGAAT 1920
RESULT 4
US-10-784-870-7
; Sequence 7, Application US/10784870
; Publication No. US20040142837A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOKI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/10/784, 870
; CURRENT FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7

; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-10-784-870-7
Alignment Scores:
Pred. No.: 7.58e-270 Length: 1923
Score: 3072.00 Matches: 596
Percent Similarity: 97.0% Conservative: 23
Best Local Similarity: 93.4% Mismatches: 19
Query Match: 94.0% Indels: 0
DB: 7 Gaps: 0
US-10-784-870-4 (1-639) x US-10-784-870-7 (1-1923)
Qy 2 ArgIleValIleValAlaPheLeuSerValIleuSerAlaAlaAlaIleLeuSerThrValAla 21
Db 7 AACGAGAAAGAAAGTGTATTTATCTGTTTATCACTGACGAGATTTGTCAGCTGTCG 66
Qy 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheIleGln 41
Db 67 TTAAATATATCATCTGACGATGGTGCAAGGAATTTGATCTGATTTCAAGGAATTCAG 126
Qy 42 ThrThrThrAspValSerGlyPheSerIleGlnArgGlnThrGlyAlaAlaIleLeu 61
Db 127 ACAACACTGATGCTAAAGGTTTCTCCAAACAGGCGACAGCTGTCGCTGCTTCTCG 186
Qy 62 ValGlnSerGlyAsnValIleLeuLeuValGlyLeuLeuValIleValPro 81
Db 187 GTGGAATCTAAATGTAATGAAATCTCCAAAGGTTTGCAAGAAAGCTTGAAACGTCGG 246
Qy 82 AlaAsnAsnIleValIleValAlaPheAsnGlyProIleLeuGlnIleThrGln 101
Db 247 GCAATATATATATCAT 306
Qy 102 LysLeuGlnIleThrThrGlyAlaValIleValIleLeuAspTrpIleProAspTrpAlaIleVal 121
Db 307 CAGCTGGAAGAAAGCAGGCGCAAGATTCGACTACATACCTGATTAATGCTTACATGTC 366
Qy 122 GlnIleGlnIleAspValIleGlnSerIleValArgSerIleGlnIleValIleGlnIleVal 141
Db 367 GAGTATGAGGCGCATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 426
Qy 142 ProIleuProIleuProIleuProIleuProIleuProIleuProIleuProIleuProIleu 161
Db 427 CTTATTTTGGCGATATACAGAAATGATCCCACTTTTCCAAAGGGGCGATACAGACTT 486
Qy 162 ValIleValAlaLeuAlaLeuAspTrpIleValIleAsnAsnIleValIleGlnLeuArgGlyIle 181
Db 487 GTAAGAGCAGTGGCGCTGATCAAGAGCAAGAAATTAAGAGGTGCAATTAAGAGGATC 546
Qy 182 GlnGlnIleAlaGlnIleValAlaIleSerAsnAspValHisTrpIleThrAlaIleProGln 201
Db 547 GAACAAATGCAAAATTCGCAATTAAGCAATGATGTCATATTTATTCGGCAAGGCGCTGAG 606
Qy 202 TyrIleValIleMetAsnAspValAlaArgGlyIleValIleValIleAspValAlaGlnIleSer 221
Db 607 TATTAAGTATGATGAT 666
Qy 222 TyrGlyLeuTrpGlyGlnIleGlnIleValAlaValAlaAspTrpGlyLeuAspTrpGly 241
Db 667 TACGGGTGATGACAGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
Qy 242 ArgAsnAspSerSerMetHisGlnAlaPheArgGlyIleValIleThrAlaLeuTrpAlaLeu 261
Db 727 CGCAATGACAGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
Qy 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerVal 281
Db 787 GAGCGACGAAATATATCCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846


```
QY 282 LeuGlyVaengIyAlaThrAenIyegIyMeraIaProGlnIaAenLeuValPheGlnSer 301
Db 847 TTAGAAACGGCTCCAAATAAGAAATGGCGCTCAGGGGAATCTAGCTTCCAAATCT 906
QY 302 IleMetAapSerSerGlyGlyLeuGlyGlyLeuProSerAenLeuGlnThrLeuPheSer 321
Db 907 ATCATGATATGCGGTGGGAGCTTGGAGGACTACCTTCGAATCTGCACAACTTATTCACG 966
QY 322 GlnAlaPheSerAenIyAlaArgIleHisThrAenSerTrpGlyAlaIaValaenGly 341
Db 967 CAAGCATACAGTGTGGTGCAGAAATCATCAAACTCTGGGGAGCAGCAGTGAATGGG 1026
QY 342 AlaTrpThrAapSerAenValaAapAapTrpAlaArgIyAaenAapMetThrIle 361
Db 1027 GCTTACACAAAGATTCAGAAATGATGATGATCTATGTGGCAAAATAATGATGACATC 1086
QY 362 LeuPheAlaIaIaGlyAaenGlyAargProAenGlyGlyThrIleSerAlaProGlyThrAla 381
Db 1087 CTTTGGCTGCGGGGAATGAAGACCGAAGCGCGGAACATCACTGACACAGCAGCAGCT 1146
QY 382 LysAaAaIleThrValGlyAlaThrGluAenLeuArgProSerPheGlySerTrpAla 401
Db 1147 AAAAATCCAAATACAGTCGAGCTACGAAACCTCCGCCCAAGCTTGGGCTTATGCG 1206
QY 402 AspAenIleAaenIyAlaIaGlnPheSerSerArgIyProThrIyAaenGlyAargIle 421
Db 1207 GACAAATTCACACATGTCGACAGCTCTCTTCACGTGACGACGACAAAGATGAGCGGATC 1266
QY 422 LysProAapValMetAlaProGlyThrTrpIleLeuSerAlaAargSerSerLeuAlaPro 441
Db 1267 AAACCGGATGTCAATGGCAGCGGGAAGCTTCACTATCAACAAAGATCTTCTTGGACCG 1326
QY 442 AapSerSerThrThrAlaAenIleAapSerIyGlyAlaTrpMetGlyGlyThrSerMet 461
Db 1327 GATTCCTCTCTGGGGGGAACATGACATGAATATGCAATACATGCTGACAGCTCATG 1386
QY 462 AlaThrProIleValIaIaGlyAaenValaIaGlnLeuArgIleHisPheValIyAaenAarg 481
Db 1387 GCTACACCAATCGTGTCTGAAACGTGGCAGACCTTCGTACAGCAATTTGTGAAAAACAG 1446
QY 482 GlyIleThrProIyAapSerSerLeuLeuValaIaLeuIleAlaGlyAlaIaAapVal 501
Db 1447 GGCATCACACCAAGCTTCTCTATTAAGCGGCACTGATTGCCGGTGCAGCTGACATC 1506
QY 502 GlyLeuGlyTrpProAenGlyAaenGlnGlyTrpGlyAargValThrIleAapIySerLeu 521
Db 1507 GGGCTTGGCTAACCCGAAACGGTAAACAAAGATGGGGAACGATGACATTGATTAATCCCTG 1566
QY 522 AenValAlaTrpValAaenGlySerSerAlaLeuSerThrSerGlnIyAaIaThrTrpYr 541
Db 1567 AAGCTTGCCATATGTAACAGATCCAGTCTCTATTCACACGCAAAAGGAGCATCTCG 1626
QY 542 PheThrAlaThrAlaGlyIyAapProLeuValIleSerIleValIlePheAapAlaProAla 561
Db 1627 TTTACTGCTACTGCGGCAAGCTTGAATAATCTCCCTGGTATGCTGATGCCCCCTGG 1686
QY 562 SerThrThrAlaSerValThrIleValaAaenAapLeuAapLeuValIleThrAlaProAen 581
Db 1687 AGACACACGCTTCCTCGTAACGCTTGCAATGATCTGACCTTGCAATCCGCTCCAAAT 1746
QY 582 GlyThrAargTrpValGlyAaenAapPheSerAlaProPheAaenAapTrpAapGlyAarg 601
Db 1747 GGCACACAGATATTAAGAAATGACTTATCTCGCCATACATATGATTAAGGATGGCCCG 1806
QY 602 AenAenValGluAaenValPheIleAaenSerProGlnSerGlyThrTrpThrIleGluVal 621
Db 1807 AATTAAGTGAATAATATATTAATTAATGACCAACAAAGCGGAGCATTAATTAAGTAAGTA 1866
QY 622 GlnAlaTrpAaenValProValGlyProGlnAaenPheSerIleLeuAlaIleValaen 639
Db 1867 CAGGCTTAATACATGACGGGTGAGCACAGAACTTCTCGTGGCAATTGGAAT 1920
```

```
RESULT 5
US-09-920-954-5
; Sequence 5, Application US/09920954
; Publication No. US20020064854A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUSHISA
; APPLICANT: KOBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-09-920-954-5

Alignment Scores:
Pred. No.: 2,676-269 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 22
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 3 Gaps: 0

US-10-764-870-4 (1-639) x US-09-920-954-5 (1-1923)

QY 2 ArgGlyGlyGlyValaPheLeuSerValIleuSerAlaIaIaIleuSerThrValaIa 21
Db 7 AAGAAAGAAAGAGTGTATTTATCTGTATATCAGCTGACGATTTTTCGACTGTTCG 66
QY 22 LeuAaenAapSerSerAlaGlyAaenAlaArgThrPheAapLeuAapPheIyGlyIleGln 41
Db 67 TTAAGTATTCATCTGCGAGGTGTGCGAAGAAATTTGATCTGGAATTTCCAAAGAAATTCG 126
QY 42 ThrThrThrAapValaSerGlyPheSerIyGlnAargGlnThrGlyAlaIaIaPheLeu 61
Db 127 ACAACAACTGATGCTAAAGGTTTCTCAAGCAGGGGCAAGACTGGTGTGCTTTCTG 186
QY 62 ValGluSerGluAaenValIyLeuLeuLeuGlyLeuLeuGlyGlyLeuGlnThrValPro 81
Db 187 GTGAATCTGAATAATGTGAACCTCCCAAGAGTTTGCAAGAAAGAACTTGAAACAGTCCG 246
QY 82 AlaAaenAenIyLeuHisIleValaGlnPheAenGlyProIleLeuGlnGlnThrIyGln 101
Db 247 GCAAAATTAATACCTCAATTTATTCAAATTCATGAGCAATTTTGAAGAAACAAACAG 306
QY 102 LysLeuGlnThrThrGlyAlaIyAaIyIleLeuAapTrpIleProAapTrpAlaTrpIleVal 121
Db 307 CAGCTGAAAGAAAGAGGGGCAAGAAAGATTCGACCTACATACGATTAATGCTTAACATTC 366
QY 122 GlyTrpGlyGlyAapValGlnSerIyAaIaAargSerIleGlnIleValaGluSerValGlu 141
Db 367 GAGTATGAGGGGAGATGTTAATGACCAACAGCAACATTAAGCAGTGGAAATCCCGTAG 426
QY 142 ProTrpLeuProIyTrpIyIleAapProGlnLeuPheThrIyGlyAlaSerThrLeu 161
```

Db	427	CCTTATTTC	CGCATATACAAATAGATCCCCAGCTTTTCACAAAGGGC	ATCAGACTT	486
Oy	162	VallYsAlaLeuAlaLeuAspThrLySGlnMetAsnLySGluValGlnLeuArgLyIle			181
Db	487	GTAAAGCAGTGGCGCTTGATACAAAGCAGAAATAAAGAGGTGCATTAAGAGGCAAC			546
Oy	182	GIUGLIUleAlGlnTyValAlaSerAsnAspValHisTyrTlleThralaLysProGlu			201
Db	547	GAACAAATCGCAAAATTCGCCAATTAAGCAATGATGTGCTAATATTAACGCAAAAGCCTGAG			606
Oy	202	TyrLyValMetAsnAspValAlaArgGlyIleValLyValAspValAlaGlnSerSer			221
Db	607	TATTAAGTGATAAAGATGTTCGGGTGAATTTGTCAAAGCCGAAATGTGGCTCAGAGCAC			666
Oy	222	TyrgLyLeuTyrgLyGlnLyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly			241
Db	667	TACGGGTGTATGAGCAAGAGCAAGATCGTAGCGGTTCCGATACAGGGCTTATACAGGT			726
Oy	242	ArgAsnAspSerSerMetHisGlnAlaPheArgGlyLysTlleThralaLeuTyralaLeu			261
Db	727	CGCAATGACAGTTCCATGCATGAACCTTCGCGGGAATAATTACGCATTAATATCATGTG			786
Oy	262	GlyArgThrAsnAsnAlaAsnAspThrArgGlyHisGlyTyrHisValAlaGlySerVal			281
Db	787	GGACCGAGCAATTAAGCCCAATGATACGAATGTGTCAATGTGACCAATGTGGCTTCGCTCGTA			846
Oy	282	LeuGlyAsnGlyAlaThrAsnLySGlyMetAlaProGlnAlaAsnLeuValPheGlnSer			301
Db	847	TTAGAAACCGGCTCCACTATTAAGAAATGGCGCTCAGGCGCAATCTAGCTTTCCAACTT			906
Oy	302	IleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSer			321
Db	907	ATCATGAGATAGCGGCTGGGGGACCTTGAGAGGACTACCTTGCATCTGCAAACCTTAATTCAGC			966
Oy	322	GlnAlaPheSerAlaGlyValaArgGlyLeHisThrAsnSerTyrGlyValAlaValaGly			341
Db	967	CAGCATACAGTGCAGTGGGCCAGAAATTCATTAACCTCGGGAGGACAGATGATGATGG			1022
Oy	342	AlaTyrThrThrAspSerArgAsnValAspAspTyrValaArgLyAsnAspMetThrIle			361
Db	1027	GCTTACACACAGATTTCCGAAATGTGATGACTATGTGCGCAAAATGATATGACATC			1088
Oy	362	LeuPheAlaAlaGlyAsnGlyLysArgProAsnGlyGlyTyrHisSerAlaProGlyThrAla			381
Db	1087	CTTTTCGTCGCGGGAATGAAGACCGAACGCGGACATCAGTGCACACAGGACAGGT			1144
Oy	382	LyAsnAlaAlaThrValaGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAla			401
Db	1147	AAAAATGCAATACAGTCCGAGCTACGGAAAACTCCGCCCAAGCTTTGGGCTTAATGCG			1206
Oy	402	AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLyAsnArgLyArgIle			421
Db	1207	GACATATATACAACTGGGCAACAGTTCTCTTACCGTGAACCGAAGAGATGAGAGGATC			1266
Oy	422	LysProAspValMetAlaProGlyTyrTyrTlleLeuSerAlaArgSerSerLeuAlaPro			441
Db	1267	AAACCGAGTGCATGCGACCGGGAACGTTCAATCATACAGCAAGATCTTCTTTGCAACG			1322
Oy	442	AspSerSerPheTyrAlaAsnHisAspSerLyTyrAlaTyrMetGlyGlyTyrHisMet			461
Db	1327	GATTCCTCTCTTCGGGCAACCATGACAGTAATATGCAATACATGGGTGGAACGTCAGT			1388
Oy	462	AlaThrProIleValAlaGlyAsnValAlaGlnLysArgGlyHisPheValLyAsnArg			481
Db	1387	GCTACACCGATCGTGTCTGGAAACGTGGACAGCTTCGAGCAATTTGTGAAAAACAGA			1444
Oy	482	GlyIleThrProLysProSerLeuLeuLyAsnAlaAlaLeuAlaGlyValaAlaAspVal			501
Db	1447	GGCAATCACACCAAACTTCTCTATTAAGAGGGGACAGATTGGCCGGTGCAGCTACATC			1506
Oy	502	GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyValArgValThrLeuAspLysSerLeu			521

```

Db      1507 GGCCTGGCTACCCGAAACGGTAACCAAGATGGGAGAGATGACATTGGTAATCCCTCG 1506
Qy      522 AenVala1a1yYrValaengJusSerAlaLeuSerThiSergJnlyPa1a1ThrYrThr 541
Db      1567 AACGTGGCTATGTGGAACGAGTCCAGTTCTCTATCCACCAAGCAAAAAGGACGTACTCG 1626
Qy      542 PheThra1a1Thra1a1aGlyVyrProLeuLy1leSerLeuValTtpSeraspAlaProAla 561
Db      1627 TTTACTGCTACGTCGGCGAAGCCCTTTGAAAATCTCCCTGGATGCTGATGCCCTCGCG 1686
Qy      562 SerThiThra1a1SerVal1ThrLeuValAsnAspLeuAspLeuVal1IleThra1a1ProAsn 581
Db      1667 AGCAACACTGCTTCGGTAACGCTTGACCAAGATCTGGACCTTGTCATTATCCGCTCCAAAT 1746
Qy      582 GlyThra1a1aGlyYrVala1a1aAsnAspPheSerAlaProPheaspAsnAsnTtpAspGlyArg 601
Db      1747 GGCACACAGTATGTAGGAATGATCTTACTTCTCCCAATACATATGATTAATCTGGAGTGGCGC 1806
Qy      602 AsnAsnVala1a1aAsnVala1a1Phe1a1aSerProGlnSerGlyThrYrThr1leGluVal 621
Db      1807 AATAACGTAAGAAATGTATTTATTAATGACACCAAAAGCGGAGCGTATTAATTAAGAGTA 1866
Qy      622 Gln1a1a1YrAsnVala1ProVala1GlyProGlnAspPheSerLeuAla1leVala1a1 639
Db      1867 CAGGCTTATTAACGTACCGGTGGACCAAGACCTTCTCGTTGGCAATTTGTGAAAT 1920

RESULT 6
US-10-456-479-3
; Sequence 3, Application US/10456479
; Publication No. US20040072321A1
; GENERAL INFORMATION:
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: TAKIMURA, YASUSHI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: NOMURA, MASAFUMI
; APPLICANT: KOBAYASHI, TOHRU
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 238700USO
; CURRENT APPLICATION NUMBER: US/10/456, 479
; CURRENT FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: JP 2002-186387
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: JP 2002-304232
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-KP43
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1920)
; OTHER INFORMATION:
US-10-456-479-3

Alignment Scores:
Pred. No.:      2,67e-269      Length:      1923
Score:          3066.00      Matches:      595
Percent Similarity: 96.9%      Conservative: 23
Best Local Similarity: 93.3%      Mismatches:  20
Query Match:    93.8%      Indels:      0
DB:             7      Gaps:      0

US-10-784-870-4 (1-639) x US-10-456-479-3 (1-1923)

Qy      2 Arg1ySylYslySValPheLeuSerValLeuSerAla1a1a1leLeuSerThra1a1a 21
Db      7 AAGAGAGAAAAGCGTTTATATCTGTTTATCAAGCTCAGCAATTTTGTGACCTGTGGC 66
Qy      22 LeuAsnAsnProSerAla1a1aSpAlaArgThrPheAspLeuAspPheLyg1y1leGln 41

```

Db 67 TTAAGTATCATCTCAGGTGGTGCAGGAATTTTGTATCTGGATTTCAAGGAATTCAG 126
Qy 42 TTTTThThApValSerGlyPheSerIySGInArgInThrGlyAlaAlaPheIeu 61
Db 127 ACAACAACGATGCTAAAGGTTTCTCCAGCAGGGGCAAGCTGGTCTGCTTTTCG 186
Qy 62 ValGluSerGlyuAsnValIySleuIeuIySGIlyLeuIeuIySleuGluThrValPro 81
Db 187 GTGGAATCTGAAGAAATGTGAATCTCCAAAAGTTTCCAGAAAGAGCTTGAACAGTCCG 246
Qy 82 AlaAspAsnIySleuIySleuAlaIyValGlnPheAsnGlyProIleuGluIyThrIyGln 101
Db 247 GCAATTAATAAATCTCCATATTAATCCAAATGAGCAATTTAGAAAGCAAAACAG 306
Qy 102 IySleuGluThrThrGlyAlaIySleuIeuIySleuPheIyProAspIyThrIleVal 121
Db 307 CAGCTGGAAAAACAGGGGCAAGATTCGACTATACCTGATTAATTCATTAATTCG 366
Qy 122 GIUTYrGluGIyAspValGlnSerIyValArgSerIleGluIySleuValGlu 141
Db 367 GAGTATGAGGGGCGATTTAACTGACACAGCAACAGCAATGAGCACTGGAAATCCGAG 426
Qy 142 ProIyLeuProIySleuIySleuIySleuPheIySleuIySleuIySleuIySleu 161
Db 427 CCTTATTTGCCATATTAAGAAATAGATCCCAAGCTTTTCAAAAAGGGGCATCAGAGCTT 486
Qy 162 ValIyValAlaValAlaIeuAspThrIySGInAsnAsnIySGIlyValGlnIeuIySleu 181
Db 487 GTPAAAGAGGTGGCTGTATACAAAGCAAAAGAAAGAGGTGAATTAAGAGGCAATC 546
Qy 182 GIUGluIleAlaGlnIyValAlaSerAspAspValIySleuIySleuIySleuIySleu 201
Db 547 GAAACAATGCAACAAATTCGCAATTAAGCAATGATGCTATTAATTCGCAAAAGCTGAG 606
Qy 202 IyIyValIySleuIySleuIySleuIySleuIySleuIySleuIySleuIySleu 221
Db 607 TATAAGGTATGAATGATGTGGCGGTGAATGTGCAAAAGCGGATGTGGCTCAGAGCAG 666
Qy 222 Iy 241
Db 667 TACGGGTGTATGACAAAGCAAGATGTCGCGTGGATACAGGGCTTGAATACAGT 726
Qy 242 ArgAsnAspSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 261
Db 727 CCCATGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
Qy 262 GIyATGThrAsnAsnAlaAsnAspThrAsnGlyIySleuIySleuIySleuIySleu 281
Db 787 GGAACGTAAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATG 846
Qy 282 LeuGIyAsnGlyAlaIySleuIySleuIySleuIySleuIySleuIySleuIySleu 301
Db 847 TTAGAAAGAGGTCTCCATTAATAAGAAATGGCGCTCAGAGCAATTCATCTTCCATCT 906
Qy 302 IleMetAspSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 321
Db 907 ATCAAGATGAGGGGTGGGGGACTTGAAGGACTACCTTCGAATCTGAAACCTTATTCAG 966
Qy 322 GlnAlaPheSerAlaGlyAlaIySleuIySleuIySleuIySleuIySleuIySleu 341
Db 967 CAAAGCATACAGTGTGGTCCAGAAATTCATACAAATCTCGGGGAGCAGCAGTGAATGG 1026
Qy 342 AlaIy 361
Db 1027 GCTTACAAACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1086
Qy 362 LeuPheAlaAlaGlyIySleuIySleuIySleuIySleuIySleuIySleuIySleu 381
Db 1087 CTTTTCGCTGGCGGAAATGAAGACCGAAGCGGAAACATCATGTCACAGCAGCAGCTT 1146
Qy 382 IySAsnAlaIleThrValGlyAlaIySleuIySleuIySleuIySleuIySleuIySleu 401
Db 1147 AAAAATGCAATAACAGTGGAGCTACGGAAGAACTTCGCGCAAGCTTGGTCTTATGGC 1206

Qy 402 AsnAsnIleAsnIyValAlaGlnPheSerSerIyProIyThrIyAspGlyIySleu 421
Db 1207 GAAATATCAACCAATGGGACAGTCTCTTACAGGAGCAAGAAAGATGACAGATC 1266
Qy 422 IySProAspValMetAlaProGlyIyThrIyIleuSerAlaArgSerSerIyPro 441
Db 1267 AAACGGATGTATGGACCGGGAACGTTCACTACATGCAAGAAATCTTCTTGACAGC 1326
Qy 442 AspSerSerPheTrpAlaAsnIyAspSerIySleuIySleuIySleuIySleuIySleu 461
Db 1327 GATTCCTCTCTGGGCGAAACCATGACAGTAATATGCAATGATGAGTGAAGCTCCAG 1386
Qy 462 AlaThrProIleValAlaGlyIyAsnValAlaGlnIeuAspGlyIySleuIySleu 481
Db 1387 GCTACACGATGCTGGGAAAGCTGGACAGCTTCGAGACATTTGTGAAGAAACAG 1446
Qy 482 GIyIleThrProIySleuIySleuIySleuIySleuIySleuIySleuIySleu 501
Db 1447 GGCATCAGACCAAGCCTTCTATTAAGGGGCACTGATGGCGGTGACAGCTGACATC 1506
Qy 502 GIyIeuGlyIyIyProAsnGlyIySleuIySleuIySleuIySleuIySleuIySleu 521
Db 1507 GGCCTTGGCTACCGCAACGCTAACAGGATGGGAGCAGTGAATGGATTAATCCCTG 1566
Qy 522 AsnValAlaIyIyValIySleuIySleuIySleuIySleuIySleuIySleuIySleu 541
Db 1567 AACGTGCTTATGTAAACAGTCCAGTCTCTATCCACAGCAAAAGCGAGTACTG 1626
Qy 542 PheThrAlaIySleuIySleuIySleuIySleuIySleuIySleuIySleuIySleu 561
Db 1627 TTTATCGTACTGCGGCAAGCTTGAAGAAATCTCCGTGATGATGATGATGATGATG 1686
Qy 562 SerThrThrAlaSerValIySleuIySleuIySleuIySleuIySleuIySleuIySleu 581
Db 1687 AGCACAGTACGCTCCGTAACGCTTCAATGATGATGATGATGATGATGATGATGATG 1746
Qy 582 GIy 601
Db 1747 GGCACAGTATGTAAGAAATGATCTTACTTGGCAATCAAGTAATGATGATGATGATG 1806
Qy 602 AsnAsnValGluAsnValPheIleAsnSerProGlnAsnPheSerIySleuIySleu 621
Db 1807 AATAAGTAAGAAATGATTTATTAATGACCAACAAAGGGGACGTATTAATGATGATG 1866
Qy 622 GlnAlaIyIyAsnValIySleuIySleuIySleuIySleuIySleuIySleuIySleu 639
Db 1867 CAGGCTTATTAAGTACGATCGGTTGGACCAAGACCTTCTGTTGGCAATGATGAT 1920
RESULT 7
US-10-784-870-5
/ Sequence 5, Application US/10784870
/ Publication No. US20040142837A1
/ GENERAL INFORMATION:
/ APPLICANT: TAKAIIWA, MIKIO
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SAKKI, KATSUHIISA
/ APPLICANT: KUBOTA, HIROMI
/ APPLICANT: HITOMI, JUN
/ APPLICANT: KAGEYAMA, YASUSHI
/ APPLICANT: SHIRATA, SHITSUM
/ APPLICANT: NOMURA, MASAFUMI
/ TITLE OF INVENTION: ALKALINE PROTEASE
/ FILE REFERENCE: 0327-0832-0PCT
/ CURRENT APPLICATION NUMBER: US/10/784, 870
/ PRIOR FILING DATE: 2004-02-24
/ PRIOR APPLICATION NUMBER: US/09/509, 814A
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: PCT/JP98/04528
/ PRIOR FILING DATE: 1998-10-07
/ PRIOR APPLICATION NUMBER: JP 9-274570
/ PRIOR FILING DATE: 1997-06-08
/ NUMBER OF SEQ ID NOS: 24

```

; SOFTWARE: Patencin version 3.0
; SEQ ID NO 5
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-10-784-870-5

Alignment Scores:
Pred. No.: 2,67e-269 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 7 Gaps: 0

US-10-784-870-4 (1-639) x US-10-784-870-5 (1-1923)

QY 2 Arglybulybelyvalpheluservalleuserlalaaleuserthrvala 21
DB :
QY 7 AAGAGAAAAAGGAGTTTATCTGTTTATCAGCTGCAGCGATTGTCAGCTGTCG 66
DB :
QY 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheGlyileGln 41
DB :
QY 42 ThrThrThrAspValSerGlyPheSerGlyleuArgGlnThrGlyalaaleu 61
DB :
QY 127 ACAACAACGATGCTAAAGGTTTCTCCAGACAGGCGCAGCTGGTCTGCTTTCTG 186
DB :
QY 62 ValGlySerGlyuAsnVallybyleuLeuGlybyleuLeuGlybyleuGlythrVal 81
DB :
QY 187 GTGGAAATCTGAAATGTAATGAAATCCCAAAAGTTTGGAGAAAGCTTGAAGACGCCG 246
DB :
QY 82 AlaAsnAsnLysLeuHisileValGlnPheAsnGlyProIleuGlnGlythrIleu 101
DB :
QY 247 GCMAATTAATTAACCTCCATTAATTAATTAATTAATTAATTAATTAATTAATTA 306
DB :
QY 102 LysleuGlnThrThrGlyalaValyleuAspTyrIleProAspTyrAlaTyrIleVal 121
DB :
QY 307 CAGCTGGAAGAAACAGGCGCAAGATTCGACGACGACGACGACGACGACGACGACG 366
DB :
QY 122 GlyTyrGlyuGlyAspValGlnSerLysValArgSerIleGlnHisValGlyuSerVal 141
DB :
QY 367 GAGTAAAGGCGCATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 426
DB :
QY 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThrLeu 161
DB :
QY 427 CCTTAATTTGCCGATACGATACGATACGATACGATACGATACGATACGATACGAT 486
DB :
QY 162 VallybAlaleuAlaleuAspThrLysGlnAsnAsnLysGlyValGlnLeuArgGlyIle 181
DB :
QY 487 GTAAGAGCAGTGCCTGTAATCAAGAGCAAGAAATTAAGAGGTCGATTAAGAGCATC 546
DB :
QY 182 GlyGlnIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaLysProGln 201
DB :
QY 547 GAACAATATCGCAATATCGCAATATCGCAATATCGCAATATCGCAATATCGCAATAT 606
DB :
QY 202 TyrLysValMetAsnAspValAlaArgGlyIleVallybAlaAspValAlaGlnIleSer 221
DB :
QY 607 TATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
DB :
QY 222 TyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
DB :
QY 667 TACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
DB :
QY 242 ArgAsnAspSerSerMetHisGlyAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeu 261
DB :
QY 727 CGCAATGACAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
DB :
QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerVal 281
```

```

DB :
QY 787 GACAGCAGCAATTAATGCAATGATACGAATGCTACGATGCGGCTGCTCCGTA 846
DB :
QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
DB :
QY 847 TTAGGAACAGGCTCCATTAATTAAGGAATGGCCCTCAGGCGAATCTAGCTTCCATCT 906
DB :
QY 302 IleMetAspSerSerGlyGlyLeuGlyGlyLeuPheSerAsnLeuGlnThrLeuPheSer 321
DB :
QY 907 ATCATGATTAAGCGGAGGAGCTTGAAGCTTCCATTCGATTCGAAACCTTAATTAAGC 966
DB :
QY 322 GlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerProGlyAlaAlaValAsnGly 341
DB :
QY 967 CAGACATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
DB :
QY 342 AlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThrIle 361
DB :
QY 1027 GCTTACACAAACAGATTCAGAAATGTGATGATCATGTGCGCAAAATGATATGACGATC 1086
DB :
QY 362 LeuPheAlaAlaGlyAsnGlyuArgProAsnGlyLysThrIleSerAlaProGlyThrAla 381
DB :
QY 1087 CTTTCCGCTGCGGAGATGAAGACCGAAGCGGGAACCATGATGACACAGCGACAGCT 1146
DB :
QY 382 LysAsnAlaIleThrValGlyAlaThrGlyuAsnLeuArgProSerPheGlySerTyrAla 401
DB :
QY 1147 AAAAATGCAATTAACAGCTCGAGCTACGAAACCTCCGCCAAGCTTTGGGCTTATGCG 1206
DB :
QY 402 AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIle 421
DB :
QY 1207 GACAAATTAACCATGATGGCAGAGTTCTCTTCAAGTGAACCAAGAGATGAGCGGATC 1266
DB :
QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
DB :
QY 1267 AAACCGAATGTCATGACCGGAGACGTTCAATCTACGAAAGATCTTCTCTGACCG 1326
DB :
QY 442 AspSerSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMetGlyLysThrMet 461
DB :
QY 1327 GATTCCTCTCTTCCGCGCAACCATGACGTAATTAATGATACATGAGTGAGACGTCATG 1386
DB :
QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlyuHisPheValLysAsnArg 481
DB :
QY 1387 GCTACACCGATCGTTCGTAACGTCGACGACGACGACGACGACGACGACGACGACG 1446
DB :
QY 482 GlyIleThrProLysProSerLeuLeuLysAlaAlaLeuIleAlaGlyAlaAlaAspVal 501
DB :
QY 1447 GGCATACACCAAGGCTTCTCATTAATAAGCGGACGCTGATCCGGTGCAGCTGCATC 1506
DB :
QY 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeu 521
DB :
QY 1507 GGCCTTGGCTAACCGAAGGTAACCAAGATGGGAGCGAGTGAATTTGATTAATCCCTG 1566
DB :
QY 522 AsnValAlaTyrValAsnGlyuSerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
DB :
QY 1567 AACGTTGCTTAATGTAACAGTCTCACTTCTATCCACGCAAGCAAAAGGAGATGATC 1626
DB :
QY 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAla 561
DB :
QY 1627 TTACTGCTACGCTCGGCAAGCCTTGAAATATCTCCCTGGTATGCTGATGCCCTGCG 1686
DB :
QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
DB :
QY 1687 AGCAACAAGCTTCCGTAACGCTTGCATGATCTGACCTTGTCAATCCGCTCCAAAT 1746
DB :
QY 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGlyArg 601
DB :
QY 1747 GGCACACATATGTAAGAAATGATCTTACTTGCCTATACAAATGATTAATGAGATGCCG 1806
DB :
QY 602 AsnAsnValGlyuAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGlyuVal 621
DB :
QY 1807 AATTAACGTGAAATATTAATTAATTAATGACCAACCAAGCGGAGCTATTAATGAGTA 1866
DB :
QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnAspSerLeuAlaIleValAsn 639
```

Db 1867 CAGGCTTAAACGTACCGGTGAGCACAGACCTTCTCGTTGGCAATTGGAAT 1920

RESULT 8
US-10-820-712A-2
Sequence 2, Application US/10820712A
Publication No. US20050026804A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okada, Mitsuyoshi
APPLICANT: Izawa, Yoshihumi
APPLICANT: Kobayashi, Tohru
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tetsuyoshi
TITLE OR INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-050
CURRENT APPLICATION NUMBER: US/10/820, 712A
PRIOR FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106708
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patent version 3.2
SEQ ID NO 2
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp. KSM-KP43
NAME/KEY: CDS
LOCATION: (1) .. (1920)
NAME/KEY: sig_peptide
LOCATION: (1) .. (618)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (619) .. ()
US-10-820-712A-2

Alignment Scores:
Pred. No.: 2, 676-269 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 8 Gaps: 0

US-10-784-870-4 (1-639) x US-10-820-712A-2 (1-1923)

QY 2 ArglyValLeuValPheLeuSerValLeuSerAlaAlaAlaLeuSerThrValAla 21
DB 7 AAGAAAGAAAGAGTGTATCTGTTTATCGAGTGCAGCGATTTTGTGCACTGTGCG 66
QY 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheGlySerIleGln 41
DB 67 TTAAGTAATCCATGCGAGGTGTCAGAGAAATTTTGATCTGATTTCAAGAAATTCAG 126
QY 42 ThrThrThrAspValSerGlyPheSerIleGlnArgGlnThrGlyValAlaAlaPheLeu 61
DB 127 ACAACACATGATGCTTAAGGTTTCTCCAGACAGGGGACAGCTGAGCTCTCTTTCTG 186
QY 62 ValGluSerGluAsnValIleLeuLeuGlyGlyLeuLeuValLeuGluThrValPro 81
DB 187 GTGGATATCGAAGAAATGTGAACCTCCCAAAAGGTTTGCAGAGAGAGCTTGAACAGTCCCG 246
QY 82 AlaAsnAsnLeuLeuHisIleValGlnPheAsnGlyProIleLeuGluGluThrIleGln 101
DB 247 GCAAAATTAATTAACCTCATTAATTCATTCATGACCAATTTGAGAGAAACAAACAG 306
QY 102 LysLeuGluThrThrGlyValAlaValIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
DB 307 CAGCTGGAAGAAAACAGGGGCAACAGATCTTCGACTACATCTGATTAAGCTTACATGTC 366
QY 122 GlyTyrGluGlyAspValGlnSerIleValArgSerIleGluHisValGluSerValGlu 141
DB 367 GAGTATGAGGGCGAGTGTAAAGTCAGCAACAGACCAATGACAGCTGGAAATCCGTGGAG 426

QY 142 ProTyrLeuProLysTyrIleAspProGlnLeuPheThrIleGlyValAspThrLeu 161
DB 427 CCTTATTTGCCGATATACAGAAATAGATCCCGAGCTTTTCACAAAGGGGCACTCAGAGCTT 486
QY 162 ValIleValAlaLeuValLeuAspThrIleGlnAsnValGluValGlnLeuArgIleIle 181
DB 487 GTAAAGCAGTGGCCCTGATACAAAGCAAAAAATTAAGAGTCCAAATTAAGAGCAATC 546
QY 182 GluGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaLysProGlu 201
DB 547 GACAAATGCAACATTCGCATTAAGCAATGATGTCATATTAATTAAGCAACGCTGAG 606
QY 202 TyrIleValIleAsnAspValAlaArgGlyIleValIleValAlaAspValAlaGlnSerSer 221
DB 607 TATAAGTGATGAAATGATGTTCGGTGGAAATTTCAAGCGGAGTGGCTCAGAGCAGC 666
QY 222 TyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
DB 667 TACGGGTGTATGACAAAGACAGATCGTACGGGTTCGATACAGGGCTTGATACAGGT 726
QY 242 ArgAsnAspSerSerMetHisGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeu 261
DB 727 CGCATATGACAGTTCGATGCAATGAACTTCCCGGGAAATTAATGCAATTAATGCAATG 786
QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerVal 281
DB 787 GACACGAGCAATTAATGCAATGATGCAATGATGATGATGATGATGATGATGATGATGATG 846
QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetLysLysProGlnAlaAsnLeuValPheGlnSer 301
DB 847 TTAGGAACGGCTCCATTAATTAAGAAATGGGGCTCAGGGCAATCTAGTCTTCAAACT 906
QY 302 IleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSer 321
DB 907 ATCATGATAGCGAGGGGGAATGAGAGCACTGCAATCTGCAAACTTATTCAGC 966
QY 322 GlnAlaPheSerAlaGlyValAlaArgIleHisThrAsnSerTrpGlyValAlaValAsnGly 341
DB 967 CAAGCATCAGTGGCTGAGGCCAGAAATTCATCAAACTCTGGGGGAGCAGTGAATGG 1026
QY 342 AlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThrIle 361
DB 1027 GCTTACACACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATG 1086
QY 362 LeuPheAlaAlaGlyAsnGlyLysProAsnGlyGlyThrIleSerAlaProGlyThrAla 381
DB 1087 CTTTTCGCTGCGGGAATGAAAGAACCGGAGAACCATCAGTGCACAGGACAGCT 1146
QY 382 LysAsnAlaIleThrValGlyValAlaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
DB 1147 AAAAATGAAATTAACGTGAGGCTAACGGAAACCTCCCGCAAGCTTTGGGTCTTATCG 1206
QY 402 AspAsnIleAsnHisValAlaGlnPheSerSerArgProIleProLysAspGlyValArgIle 421
DB 1207 GACATATCAACATGTGSCACAGTCTCTTCAAGTGCAGCAACAAAGATGACAGGATC 1266
QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
DB 1267 AAACCGAGATGCATGGACCGGGAAAGCTTATCATCATCAGCAAGATCTTCTTGGACCG 1326
QY 442 AspSerSerPheTrpAlaAsnHisAspSerIleGlyTyrMetGlyGlyThrSerMet 461
DB 1327 GATTCCTCTCTTGGGGCAACAGACAGTAAATATGATCATCAGGTGGGAAACGTCCATG 1386
QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHisAspValLysAsnArg 481
DB 1387 GCTACACCGATCGTTGCTGGAACGTGSCACAGCTTGTGAGCAATTTGTGAAAAACGA 1446
QY 482 GlyIleThrProLysProSerLeuLeuLysAlaIleLeuIleAlaGlyAlaAlaAspVal 501
DB 1447 GGCATCACACCAAGACCTTCTCTATTAAGCGGACAGATGCTGGGAGCTGACATC 1506

QY 502 G1YLeuGIYTYrProaenGIYAasnGIYTrpGIYArGIYThrLeuApbLYSerLeu 521
DB 1507 GGCTTGGCTAACCGAAACGGTACCAAGCATGGGACAGATGACATTAAGTAATCCCTG 1566
QY 522 AasnValATrYValaenGIUSerSerAlaLeuSerThrSerGIYValaArThrTYrThr 541
DB 1567 AACGTTCCTATGTAAGCATGAGTCCAGTCTCTATCCACAGCCAAAAGAGAGTACTCG 1626
QY 542 PhetrThrAlaThrAlaGIYLYrProLeuLYlSerLeuValTrpSerApbAlaProAla 561
DB 1627 TTTACTGCTACTGCGGCAAGCCTTTGAAATCTCCCTGATGCTGATGCTGCTGCG 1686
QY 562 SerThrThrAlaSerValTrpLeuValaAasnApbLeuValIleThrAlaProaen 581
DB 1687 AGCAACACTGCTCCGTAACGCTTGTCAATGATCTGACCTTGTCAATACCGCTCCAAAT 1746
QY 582 G1YThrArGIYrYValGIYAasnApbPheSerAlaProPheAasnAenTrpApbGIYArG 601
DB 1747 GGCACACAGTATGAGAAATGACTTACTTCGCCATACAAATGATACTGGAGTGGCCG 1806
QY 602 AasnAenValGIUaenValPheIleAenSerProGIUSerGIYThrTYrThrIleGIUVal 621
DB 1807 AATAACGTAGAAATGATATTAATTAATGACACCAAAAGCGGACGTATACAAATTGAGTA 1866
QY 622 G1nAlATrYrAenValProValGIYrProGIUaenPheSerLeuValIleValaen 639
DB 1867 CAGGCTTAATACGTACCGGTTGACACACAGACCTTCTGTTGGCAATTGTGAAT 1920
RESULT 9
US-10-820-714A-2
Sequence 2, Application US/10820714A
Publication No. US20050214922A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuyoshi
APPLICANT: Kobayashi, Noboru
APPLICANT: Sumitomo, Nobuyuki
APPLICANT: Takimura, Yasushi
APPLICANT: Sato, Teiyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251697USO
CURRENT APPLICATION NUMBER: US/10/820,714A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106709
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp. KSM-KP43
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: (1)..(618)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (619)..()
US-10-820-714A-2
Alignment Scores:
Pred. No.: 2,67e-269 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 9 Gaps: 0
US-10-784-870-4 (1-639) x US-10-820-714A-2 (1-1923)
QY 2 ArgLYbLYbLYbValPheLeuSerValLeuSerAlaAlaIleLeuSerThrValAla 21

DB 7 AAGAAGAAAAGAGTGTTTTATCTGTTTATCTGCTGACCGCATTTGTTCAGCTGTTCG 66
QY 22 LeuAasnApbProSerAlaGIYAapAlaArGIThrPheAasnApbPheLYGIYleGI 41
DB 67 TTAAGTAATCATCTGAGGTGGTGAAGGAATTTGATCTGGATTTCAAGGAATTCAG 126
QY 42 ThrThrThrAenValSerGIYrPheSerLYrGIaArGIThrGIYAlaAlaIlePheLeu 61
DB 127 ACAACAACCTGATCTAAAGGTTTCTCCAAAGCAGGAGCATGCTGTGCTGCTTTCTG 186
QY 62 ValGIUSerGIUaenValLYrLeuLeuLYrGIYLeuLeuLYrSerLeuTrpValPro 81
DB 187 GTGAAATCTGAAAATGTGAACCTCCAAAGGTTTGAGAGAGCTTGAACAGTCCG 246
QY 82 AlaAasnAenLYrLeuValIleValGIaPheAenGIYProIleLeuGIUTrpLYrGI 101
DB 247 GCAAAATATACTATCCATATTAATCAATTCAAATGACCAATTTTGAAGAAACAAACAG 306
QY 102 LYrLeuGIUThrThrGIYAlaLYrIleLeuApbTYrIleProApbTYrAlaTYrIleVal 121
DB 307 CAGCTGAAAACACAGGAGCAAGATTCGACTACATACCTGATTAATGCTTACATTTGC 366
QY 122 G1UTrpGIUaenValGIUSerLYrValArGIYrGIUaenSerIleGIUaenValGIU 141
DB 367 GAGTATGAGGCGCATGTAAAGTCAACCAAGCAACCAATGAGCACTGGAAATCCGTGAG 426
QY 142 ProTYrLeuProLYrLYrValIleAenProGIUaenPheThrLYrGIYAlaSerThrLeu 161
DB 427 CTTATTTGCCGATATACAAATAGATCCCACTTTTCAAAAAGGGCATCGAGCTT 486
QY 162 ValLYrAlaLeuValaLeuApbThrLYrGIaAasnAenLYrGIUValGIUaenArGIYIle 181
DB 487 GTAAAGACAGTGGCGCTGTATCAAAAGCAGAAAATAAAGAGGTGCAATTAAGAGCATC 546
QY 182 GIUaenValaGIUaenValaAserAasnAenValaIleTYrIleThrAlaLYrProGIU 201
DB 547 GAACAATGCAACAATTCGCAATACAAATGATGTCTATATATACGGCAAAAGCCGTAG 606
QY 202 TYrLYrValaIleAasnAenValaAaArGIYIleValLYrAlaAenValaAaGIUSer 221
DB 607 TATTAAGTATGATATATATGTTGGCTGCAATTTGTCAAAAGCGATGTGCTCGAGACGC 666
QY 222 TYrGIYLeuTYrGIYGIaenGIYIleValaIleAaenThrGIYLeuApbThrGIY 241
DB 667 TACGGGTGTATGACAAAGACAGATCGTAGCGGTGCGCATACAGGCTTGATACAGGT 726
QY 242 ArgAasnApbSerSerThrIleGIUaIlePheArGIYLYrIleThrAlaLeuTYrAlaLeu 261
DB 727 CGCAATGACAGTTCGATGATGATGAAGCTTCCGCGGAAAATTAATCATTAATATGCAATG 786
QY 262 G1YArGIThrAasnAenValaAenAenThrAenGIYHleGIYThrHleValaIleAGIYSerVal 281
DB 787 GAGCGACGAATTAATCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
QY 282 LeuGIYAenGIYAlaThrAenLYrGIYMetAlaProGIAlaAasnLeuValPheGIUSer 301
DB 847 TTAGAAAACGGCTCCACTATTAAGGAATGGCGCTCAAGCGCAATCTTACATTC 966
QY 302 IleMetAenSerSerGIYGIYLeuGIYGIYLeuProSerAenLeuGIUThrLeuPheSer 321
DB 907 ATCATGTGATAGCGGTGGGAGCTTGGAGGACTTCAATCTGCAATCTTAACTTATTCACAC 966
QY 322 G1nAlaPheSerAlaGIYAlaArGIYIleHleThrAenSerTrpGIYAlaAlaValaenGIY 341
DB 967 CAAGCATACAGTCTGCTGCTGCAAGTATTCATCAAACTCTGGGAGACAGATGAATGAG 1026
QY 342 AlaTYrThrThrAenSerArGIaenValaAenAenTYrValaArGIYLeuAenAenPheThrIle 361
DB 1027 GCTTACACAACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
QY 362 LeuPheAlaAlaGIYAenGIUaenProaenGIYGIYThrIleSerAlaProGIYThrAla 381

Db 1087 CTTTCGCTTCGCGGAATGAAGGACCGAAGCGGGAACATCATGTCACCGACAGCT 1146
Qy 382 LysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
Db 1147 AAAAAAGCAATTAACAGTCGAGCTACGGAAAACTCCGCCCAAGCTTTGGGCTTATAGCG 1206
Qy 402 AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrIleAspGlyArgGly 421
Db 1207 GACAAATATCAACATGTGGGACAGCTTCTTTCACGTGGGACCGAAGAAAGATGACCGATC 1266
Qy 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
Db 1267 AAACCGAGATGTCATGGCACCGGGAAGCTTCACTATCAGCAAGATCTTCTTGGACCG 1326
Qy 442 ArgSerSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMetGlyGlyThrSerMet 461
Db 1327 GATTCTCTCTTCGGGCGAAGCAATGACATGAAATATGATACATACATGGGTGGAACGTCCAG 1386
Qy 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlyHisPheValLysAsnArg 481
Db 1387 GCTACACCAATCTTCTCGAAACGTGGCACAGCTTCGTGAGCATTTTGTGAAAAACAA 1446
Qy 482 GlyIleThrProLysProSerLeuLeuLysAlaIleLeuIleAlaGlyAlaAlaAspVal 501
Db 1447 GGCATCACACCAAGCCTTCTCTATTAAGCGGCACTGATTCGCGTGCAGCTGACATC 1506
Qy 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrThrGlyArgValThrLeuAspLysSerLeu 521
Db 1507 GGCCTTGGCTACCCGAAACGGTACCAAGAGTGGGAGCATGATGATTAATCCCTG 1566
Qy 522 AsnValAlaIleTyrValAsnGlySerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
Db 1567 AACGTGCTTAATGTGAACAGATCCAGTCTCTATTCACACGCAAAACGACGATCTCG 1626
Qy 542 PheThrAlaIleThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAla 561
Db 1627 TTATCTGCTACTGCGCGGCAAGCCTTGAAGAACTCCCTGTATGCTGATCCCTCGG 1686
Qy 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCACACACTGCTTCCGTACCGCTTGCATGATCTGGACCTTGCTTAACCTCCCAAT 1746
Qy 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGlyArg 601
Db 1747 GGCACACAGTATGTAGAAATGACTTACTTCCGCAACATGATGATACGCGATGCGCG 1806
Qy 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
Db 1807 AATTAACGTAGAAATGTATTATTAATGACACCAAGCGGACGATATCAATTGAGTGA 1866
Qy 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGCTTATTAACGTACCGGTTGGACCAAGACCTTCTCTTGGCAATTGTGAAT 1920
RESULT 10
US-10-820-712A-11
; Sequence 11, Application US/10820712A
; Publication No. US20050026804A1
GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuoshi
; APPLICANT: Izawa, Yoshihumi
; APPLICANT: Kobayashi, Toru
; APPLICANT: Koyama, Shingo
; APPLICANT: Sato, Tatsuoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: P51701-USO
; CURRENT APPLICATION NUMBER: US/10/820,712A
; PRIOR APPLICATION NUMBER: 2004-04-09
; PRIOR FILING DATE: 2003-106708
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 11
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-KP9860
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: protease KP9860
US-10-820-712A-11
Alignment Scores:
Pred. No.: 6,696-194 Length: 1302
Score: 2237.00 Matches: 433
Percent Similarity: 99.88 Conservative: 0
Best Local Similarity: 99.84 Mismatches: 1
Query Match: 68.54 Indels: 0
DB: 8 Gaps: 0
US-10-784-870-4 (1-639) x US-10-820-712A-11 (1-1302)
Qy 206 AsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSerTyrGlyLeuTyr 225
Db 1 AATGATGTGGCCAGAGTATGTGAAACCGAGATGGCACAGACAGCTACGGTTGTAT 60
Qy 226 GlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArgAspAspSer 245
Db 61 GGCACAGGCGCATTTGTGCGAGTTGCGATCTGGATTTGGATACGAGAAACGACAGT 120
Qy 246 SerMetHisGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeuGlyArgThrAsn 265
Db 121 TCGATGCAATGAAGCTTCGCGGTAAATTAACACACTATATGCACTGGGTGGACGAAAT 180
Qy 266 AsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGlyAsnGly 285
Db 181 AATGCAATGATACGAACGATCATGTGATCCCATGTGGCAGGTTGGTATTAAGAAATGCG 240
Qy 286 AlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSer 305
Db 241 GCACGAAATTAAGAAATGCGACCTCGAACGAAATCTGGTTTTCATCATCATGATGATGC 300
Qy 306 SerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSer 325
Db 301 AGTGGTGGGCTTGGAGGCTTCCCTTCAATCTGCAAACTTATTCAAGCAAGCATTCAGT 360
Qy 326 AlaGlyAlaAspGlyIleHisThrAsnSerTyrGlyAlaAlaValAsnGlyAlaTyrThrThr 345
Db 361 GCAGGTGCGAATTCATTAACAACTCTGGGGGCGACGCGGATGGGCGCTTACACGACA 420
Qy 346 AspSerArgAsnValAlaAspTyrValArgLysAsnAspMetThrIleLeuPheAlaAla 365
Db 421 GATTCCGAATATGTGATGATCTATGAGAAATATATATGATGATTTCTTTCCGGCT 480
Qy 366 GlyAsnGlyArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaLysAsnAlaIle 385
Db 481 GGGAAATGAAGGCGGAACGGGCGTACCATCATGCTCGGTACGGCTTAATAAGCCATC 540
Qy 386 ThrValGlyAlaThrGluAsnLeuArgProSerPheIleSerTyrAlaAspAsnIleAsn 405
Db 541 ACTGTGGCGCAACCGAAACCTGGCTCAAGCTTGGTTCTTAATGCAATTAATTAAC 600
Qy 406 HisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIleLysProAspVal 425
Db 601 CAGCTTGACAGTCTTCTTCCCGTGGCCGACAAATAATGGCGAATCAAGCTGATGTC 660
Qy 426 MetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerPhe 445
Db 661 ATGGCGCCAGGCAATCATTTATCAGCAAGATCTTCTTCCACCGCATTCCTCTTC 720
Qy 446 TrpAlaAsnHisAspSerLysTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIle 465
Db 721 TGGCGAATCATGACCAAAATATGCTTATATGGGTGAACGTCATGGGCAACCGCATT 780
Qy 466 ValAlaGlyAsnValAlaGlnLeuArgLysIlePheValLysAsnArgGlyIleThrPro 485

Db 781 GTGGGGGAAATGTGCACAGCTCCGTGAGCATTTTGTGAAAATAGAGGATCACTCCT 840
Qy 486 LysProSerLeuLeuLysAlaAlaLeuLleAlaGlyAlaAlaAspValGlyLeuGlyTyr 505
Db 841 AAGCTTCCCTATTTGAAAGCAGCTTTGATTTGACAGTCTCTGTATGTTGATGGCTTAT 900
Qy 506 ProAsnGlyAsnGlnGlyTTPGlyArgValThrLeuAspLysSerLeuAsnValAlaTyr 525
Db 901 CCGAACGAAACCAAGATGGGCGGAGTGAACCTGTGATATGTTGAACGTGGCTAT 960
Qy 526 ValAsnGluSerSerAlaLeuSerThrSergLysAlaThrTyrThrPheThrAlaThr 545
Db 961 GTGAACGAATCCAGTGCCTTATCAACTAGCCAAAACGACATTACTTACTTACGCAAG 1020
Qy 546 AlaGlyLysProLeuLysLleSerLeuValTTPSerAspAlaProAlaSerThrAla 565
Db 1021 GCGGGCAAGCCATTGAAATCTCCCTGTATGTCGATGCCCTCGAACGACTACTGCT 1080
Qy 566 SerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyThrArgTyr 585
Db 1081 TCTGTAACTCTGTGATGATTTGGATTTGGTCACTTAAGACCAACGAAACGAATAT 1140
Qy 586 ValGlyAsnAspPheSerAlaProPheAspAsnAsnTTPAspGlyArgAsnAsnValGlu 605
Db 1141 GTGGGGAATGACTTCTCAGACCACTTGAACATTAAGGATGGGCCCAATTAAGTGA 1200
Qy 606 AsnValPheIleAsnSerProGlnSergLysThrTyrIleGluValGlnAlaTyrAsn 625
Db 1201 AAGTATTTATTTATTCGCCCAAGGAAACATATATCAATGAGTGCAAGCATATAT 1260
Qy 626 ValProValGlyProGlnAsnAspSerLeuAlaIleValAsn 639
Db 1261 GTCCCGGTGGACCAAAACCTCTCGTGCATTTGTGAAC 1302

RESULT 11
US-10-820-714A-12
Sequence 12, Application US/10820714A
Publication No. US20050214922A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuyoshi
APPLICANT: Kobayashi, Tohru
APPLICANT: Sumitomo, Nobuyuki
APPLICANT: Takimura, Yasuhiro
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251697US0
CURRENT APPLICATION NUMBER: US/10/820,714A
PRIOR FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106709
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.2
SEQ ID NO 12
LENGTH: 1302
TYPE: DNA
ORGANISM: Bacillus sp. KSM-kp9860
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1302)
OTHER INFORMATION: protease KP9860
US-10-820-714A-12

Alignment Scores:

Pred. No.: 6,69e-194 Length: 1302
Score: 2237.00 Matches: 433
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 68.5% Indels: 0
DB: 9 Gaps: 0

US-10-784-870-4 (1-639) x US-10-820-714A-12 (1-1302)

Qy 206 AsnAspValAlaArgGlyIleValLysValAspValAlaGlnSerSerTyrGlyLeuTyr 225
Db 1 AATGATGTGGCCAGAGGTATTTGAAAGCCGAATGTGCACAGACGACTACGGTTGTAT 60
Qy 226 GlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSer 245
Db 61 GGAACAGGCCAGATTGTCCGAGTTGCCGATCTCGATTTGATTCAGGAAGAAACGACGT 120
Qy 246 SerMetHisGlyIleAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeuGlyArgThrAsn 265
Db 121 TCCATCATATGAACTTCCCGCGGTAAATATACAGCACTATATCACTGCGGTCCGACGAAT 180
Qy 266 AsnAlaAsnAspThrArgGlyHisGlyThrHisValAlaGlySerValLeuGlyAsnGly 285
Db 181 AATGCCAAATGATCGAAGCGTCATGTATCCCAATGTGCAGATTCGGTATTAAGAAATGAGC 240
Qy 286 AlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSer 305
Db 241 GCACCAATTAAGGAATGGCACTCAAGCGAATCTGCTTTTCATTCATCAATGATAGC 300
Qy 306 SerGlyGlyLeuGlyLysLeuProSerAsnLeuGlnThrLeuPheSergLysAlaPheSer 325
Db 301 AGTGTGGGCTTGAGAGCTTGCTTCAATCTGCAAACTTATTCAGCCAAAGCATTCAGT 360
Qy 326 AlaGlyValArgIleHisThrAsnSerTTPGlyValAlaValAlaGlnGlyAlaTyrThr 345
Db 361 GCAGGTGCCGAATTAATCAAACTCTGGGGGCGAGCGGTGAATGGGCTTCACACACA 420
Qy 346 AspSerArgAsnValAlaAspPyrValArgLysAsnAspMetThrIleLeuPheAlaAla 365
Db 421 GATTCCGAATATGATGATGATCTATGAGGAAGAAATGATATGCGATTCTTTTCGCGGCT 480
Qy 366 GlyAsnGluArgProAsnGlyLysThrIleSerAlaProGlyThrAlaLysAsnAlaIle 385
Db 481 GGGAAATGAAGGCGCGAAGCGCGTACATCATGACGTCGCTGTAAAGCGCATC 540
Qy 386 ThrValGlyValAlaThrGluAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsn 405
Db 541 ACTGTGGGCGCAACCAAAACCTGCGTCCAAAGCTTCGTATCAAGTAATATTAATAC 600
Qy 406 HisValAlaGlnPheSerSerArgLysProThrLysAspGlyArgIleLysProAspVal 425
Db 601 CAGCTTGACAGTCTCTTCCCTGCGGCCGACAAAGATGGCGAATCAAGCTGATGTC 660
Qy 426 MetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerPhe 445
Db 661 ATGGCCCGAGGACATACATTTTATCAGCAAGATCTTCTTCCGACCCGATTCCTCCTTC 720
Qy 446 ThrAlaAsnHisAspSerLysTyrAlaIleMetGlyGlyThrSerMetAlaThrProIle 465
Db 721 TGGGCGAATCATGACGAATATGCTATATGGTGGAAACGTCCATGGCAACCGGAT 780
Qy 466 ValAlaGlyAsnValAlaGlnLeuArgGlyHisAlaPheValLysAsnArgGlyIleThrPro 485
Db 781 GTTGGGGGAATGTGCACAGCTCCGTGAGCATTTGTGAAAATAGAGGAATCACTCCT 840
Qy 486 LysProSerLeuLeuLysAlaAlaLeuLleAlaGlyAlaAlaAspValGlyLeuGlyTyr 505
Db 841 AAGCTTCCCTATTTGAAAGCAGCTTTGATTTGACAGTCTCTGTATGTTGATGGCTTAT 900
Qy 506 ProAsnGlyAsnGlnGlyTTPGlyArgValThrLeuAspLysSerLeuAsnValAlaTyr 525
Db 901 CCGAACGAAACCAAGATGGGCGGAGTGAACCTGTGATATGTTGAACGTGGCTAT 960
Qy 526 ValAsnGluSerSerAlaLeuSerThrSergLysAlaThrTyrThrPheThrAlaThr 545
Db 961 GTGAACGAATCCAGTGCCTTATCAACTAGCCAAAACGACATTACTTACTTACGCAAG 1020
Qy 546 AlaGlyLysProLeuLysLleSerLeuValTTPSerAspAlaProAlaSerThrAla 565
Db 1021 GCGGGCAAGCCATTGAAATCTCCCTGTATGTCGATGCCCTCGAACGACTACTGCT 1080

QY 566 SerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyThrArgTyr 585
Db 1081 TCGTAACCTCGGTGATGATTTGGATTGTGATTCATTCACACCAAGCAAGATAT 1140
QY 586 ValGlyAsnAspPheSerAlaProPheAspAsnTrpAspGlyArgAsnValGlu 605
Db 1141 GTGGGAAATGACTTCAGACCATTTGACATTAACGTGGATGGCCGCAATACGTAGAA 1200
QY 606 AsnValPheIleAsnSerProGlnSerGlyThrThrIleGluValGlnAlaTyrAsn 625
Db 1201 AATGATTTATTAATTCGCCCCAAGTGAACATTAACATTAAGGTGCAAGCATATTAAT 1260
QY 626 ValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1261 GTCCCGGTGGACCAAAAACCTTCGTGGCAATTGTGAC 1302

RESULT 12
US-10-820-712A-13
Sequence 13, Application US/10820712A
Publication No. US20050026804A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuyoshi
APPLICANT: Izawa, Yoshitumi
APPLICANT: Kobayashi, Tokuo
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tetsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-USO
CURRENT APPLICATION NUMBER: US/10/820,712A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106708
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.2
SEQ ID NO 13
LENGTH: 1302
TYPE: DNA
ORGANISM: Bacillus sp. KSM-9865
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1302)
OTHER INFORMATION: Protease 9865
US-10-820-712A-13

Alignment Scores:
Pred. No.: 1,986-189 Length: 1302
Score: 2188.00 Matches: 419
Percent Similarity: 99.34 Conserved: 12
Best Local Similarity: 96.54 Mismatches: 3
Query Match: 67.04 Indels: 0
DB: 8 Gaps: 0

US-10-784-870-4 (1-639) x US-10-820-712A-13 (1-1302)

QY 206 AsnAspValAlaArgGlyIleValIleValAspValAlaGlnSerSerTyrGlyLeuTyr 225
Db 1 AATGATGTTGACGCGGAAATTCGCAAGCGATGTGGCGCAGACGCTAOCGGTTGAT 60
QY 226 GlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSer 245
Db 61 GACACAGACAGATGTAGCGGTTCGCGATACAGGCGCTTGATACAGGTGCGCATGACAGT 120
QY 246 SerMetHisGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeuGlyArgThrAsn 265
Db 121 TCGATGCAATGAAGCTTCGCGGAGAAATTAATGCAATTAATGCAATGCGACGCAAT 180
QY 266 AsnAlaAsnAspThrAsnGlyArgIleThrIleValAlaGlySerValLeuGlyAsnGly 285
Db 181 AATGCCAATGATACGAATGTGTCATGTCATGCGATGGCGGTCCGTATTGGAACCGC 240
QY 286 AlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSer 305
:::|||||

Db 241 TCCACTAATTAAGAAATGCGCCTCAGCGGAATCTAGTCTTCAATCTATCATGATAC 300
QY 306 SerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSer 325
Db 301 GGTGGGGGAGCTTGGAGAGCACTACCTTCGAAATGCAAACTTATTCAGCCAGCAATACGT 360
QY 326 AlaGlyAlaArgIleHisThrAsnSerTrpGlyAlaIleValAsnGlyAlaTyrThrThr 345
Db 361 GCTGGTCCAGAAATTCATCAAACTTCGGGAGAGCAAGTGAATGGGGCTTACACACA 420
QY 346 AspSerArgAsnValAspAspTyrValArgLysAsnAspMetThrIleLeuPheAlaIle 365
Db 421 GATTCAGAAATGTGATGATGATGTGGCCAAATGATTAACATCTTTGCTGCTCC 480
QY 366 GlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaLysAsnAlaIle 385
Db 481 GGGAAATGAAGACCGAAGCGCGAAACCATCATGTGACACAGACACAGCTAAAAAATGCAATA 540
QY 386 ThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsn 405
Db 541 ACAATCGAGCTACGAAAACTCCGCCAAGCTTCGGGTCTTATGGCGACATATTAAC 600
QY 406 HisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIleLysProAspVal 425
Db 601 CATGTGGCAGTTCTCTTCACTGACCGACCAAAAGATGACCGATCAAAACCGATGTC 660
QY 426 MetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerPhe 445
Db 661 ATGGCACCGGAAACGTTCTATCATCATGACAAAGATCTTCTGTGACCGGATTCCTCCTTC 720
QY 446 TrpAlaAsnHisAspSerLysTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIle 465
Db 721 TGGCGAACCATGACAGTAATATGCAATCATGGGTGGAAAGTCATGAGCTACACCGATC 780
QY 466 ValAlaGlyAsnValAlaGlnLeuArgGlyLysIlePheValLysAsnArgGlyIleThrPro 485
Db 781 GTTGCTGGAACGTTGACAGCTTCGTGAGCATTTTGGAAACAGAGGATCAACCA 840
QY 486 LysProSerLeuLeuValAlaLeuIleAlaGlyAlaAlaAspValGlyLeuGlyTyr 505
Db 841 AAGCTTCTCTATTAAGACGCGACATGATCCCGGTGACGACGACATCGGCTTGGCTAC 900
QY 506 ProAsnGlyAsnGlnGlyTrpGlyArgValThrLeuAspLysSerLeuAsnValAlaTyr 525
Db 901 CCGAACGTTACCAAGATGGGAGCGAGTACATTTGATTAATCCTTAACGTTGCTAT 960
QY 526 ValAsnGlnSerSerAlaLeuSerThrSerGlnValAlaThrTyrThrPheThrAlaThr 545
Db 961 GTGAACGAGTCAAGTCTCTATTCACACGCAAAAGCGAGTACTCGTTACTGCTACT 1020
QY 546 AlaGlyLysProLeuLysIleSerLeuValTrpSerAspAlaProAlaSerThrThrAla 565
Db 1021 GCCGCAAGCTTTGAAGATCTCCCTGTATGTCTATGCTCCCTGAGACACAACGCT 1080
QY 566 SerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyThrArgTyr 585
Db 1081 TCGTAACCTCGGTGATGATTTGGATTGTGATTCATTCACACCAAGCAAGATAT 1140
QY 586 ValGlyAsnAspPheSerAlaProPheAspAsnTrpAspGlyArgAsnValGlu 605
Db 1141 GTGGGAAATGACTTCAGACCATTTGACATTAACGTGGATGGCCGCAATACGTAGAA 1200
QY 606 AsnValPheIleAsnSerProGlnSerGlyThrThrIleGluValGlnAlaTyrAsn 625
Db 1201 AATGATTTATTAATTCGCCCCAAGTGAACATTAACATTAAGGTGCAAGCATATTAAT 1260
QY 626 ValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1261 GTCCCGGTGGACCAAAAACCTTCGTGGCAATTGTGAC 1302

RESULT 13
US-10-820-714A-14
Sequence 14, Application US/10820714A

```
/ Publication No. US20050214922A1
/ GENERAL INFORMATION:
/ APPLICANT: KAO CORPORATION
/ APPLICANT: Okuda, Mitsuyoshi
/ APPLICANT: Kobayashi, Tohru
/ APPLICANT: Sumitomo, Nobuyuki
/ APPLICANT: Takimura, Yasuichi
/ APPLICANT: Sato, Tsuyoshi
/ TITLE OF INVENTION: ALKALINE PROTEASE
/ FILE REFERENCE: 251697USO
/ CURRENT APPLICATION NUMBER: US/10/820,714A
/ PRIOR FILING DATE: 2004-04-09
/ PRIOR APPLICATION NUMBER: 2003-106709
/ PRIOR FILING DATE: 2003-04-10
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 14
/ LENGTH: 1302
/ TYPE: DNA
/ ORGANISM: Bacillus sp. KSM-9865
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1302)
/ OTHER INFORMATION: Procease 9865
US-10-820-714A-14

Alignment Scores:
Pred. No.: 1,98e-189 Length: 1302
Score: 2188.00 Matches: 419
Percent Similarity: 99.3% Conservative: 12
Best Local Similarity: 96.5% Mismatches: 3
Query Match: 67.0% Indels: 0
DB: Gaps: 9

US-10-784-870-4 (1-639) x US-10-820-714A-14 (1-1302)

QY 206 AenApVa1AaArg1y1eVal1yVal1AaApVa1AaG1nSerSerTyrg1yLeuTy 225
DB 1 AATGATGTTTCACTGGAATTTGTCMAAGCGGATGTGGCGCAAGCACTTCGGGTTGTAT 60
QY 226 GlyGlnGlyGln1eVal1AaVal1AaApThG1yLeuApThrg1yArGaenApSer 245
DB 61 GGAACAAGACAGATCGAAGCGGTTGCGGATACAGGGCTTGATACAGGTGCAATGACAGT 120
QY 246 SerMetH1eG1n1AaPheArG1yLyP1eThr1AaLeuTyra1AaLeuG1yArGThraPn 265
DB 121 TCGATGCAATGAACCTTCCGGGGAAATTAATGCAATTAATGCAATTCATTCGACGACGAAAT 180
QY 266 AenA1AaenApThraenG1yH1eG1yThrH1eVal1AaG1ySerVal1eug1yAenG1y 285
DB 181 AATGCCAAATGATTCGAATGATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 286 A1aThrAen1yG1yMetA1aProG1n1Aaen1eVal1PheG1nSer11eMetAaPser 305
DB 241 TCCACTAAATGAAGAAATGGGCTCAAGCGGAATCTAGTCTTCAATCTATACATGATAGC 300
QY 306 SerG1yG1yLeuG1yG1yLeuP1eSerA1eLeuG1nTh1eupHeserG1n1AaPheSer 325
DB 301 GGTGGGGGACTTGGAGACTTACCTTCGAATCTCAACCTTATTCAGCCAAAGCATACAGT 360
QY 326 A1aG1yA1aArg11eH1eThraenSerTPG1yA1aA1eVal1AaenG1yA1yThrTrHr 345
DB 361 GCTGGTGCAGAAATTCATACAACTCTCTGGGGAGAGAGCATGATGGGCTTACACAAAC 420
QY 346 AspSerArgAaenVal1AaPArTyra1yArG1yAaenApMetThr11eLeuPheA1a1a 365
DB 421 GATTCAGAAATGTGATGATCTATGTGCGCAAAATGATATGACATCTTTTGGTGGC 480
QY 366 GlyH1eG1yA1yProAenG1yG1yThr11eSerA1eProG1yThra1yAaenA1a1e 385
DB 481 GGGAAATGAAGACCGAAGCGGAAACATCATGAGCAACAGGCAAGCTAAATTAATGCAATA 540
QY 386 ThrValG1yA1aThrg1uAen1eA1yGProSerPheG1ySerTyra1AaPheAen11eAa 405
```

```
DB 541 ACAAGTGGAGCTACGAAAACTCCGCCCAAGCTTGGGCTTATGCGGCAATATAC 600
QY 406 H1eVal1AaG1nHeserSerArG1yProTh1yAaP1yA1yG1yLeuP1eAaPva1 425
DB 601 CATGTGCAAGATTTCTTCACTGAGCCGACAAAGATGAGCGATCAACAGGATGTC 660
QY 426 MetA1aProG1yThrTyra1eLeuSerA1aArgSerSer1eA1aProAaPserPhe 445
DB 661 ATGGCAAGGGAAGCTTCAATACATACGAAGATCTTCTTTCACACCGAATTCCTTC 720
QY 446 TPra1AaenH1eAaPserTyra1yArG1yMetG1yG1yThrSerMetA1aThPro11e 465
DB 721 TGGGCAACCATACAGTAATATGATACATAGGTGGAGAGTTCATGCTTACACGATC 780
QY 466 Val1AaG1yAaenVal1AaG1n1eA1yA1yA1yA1yA1yA1yA1yA1yA1yA1yA1y 485
DB 781 GTTGTGAAACGTGGACAGCTTCTGAGCAATTTGTGAAAACAGAGGCATCACCA 840
QY 486 LyPProSer1eLeuTyra1Aa1eA1y1eA1yA1yA1yA1yA1yA1yA1yA1yA1yA1y 505
DB 841 AAGCTTCTCTATTAAGGCGCATGATTCGCGGACAGACATCGGCTTGGCTAC 900
QY 506 ProAenG1yAaenG1yTPG1yArG1yA1yTh1eAaP1ySer1eAaenVal1A1yTy 525
DB 901 CCGAACGTAACCAAGATGGGACAGATGACATTCGATTAATCCTTAACGTTGCTAT 960
QY 526 Val1AaenG1ySer1eA1eA1ySerTh1yA1yA1yA1yA1yA1yA1yA1yA1yA1yA1y 545
DB 961 GTGAACAGATCGATCTCTATTCACAGCAAAAGCGACGATCGTTTACGTGACT 1020
QY 546 A1aG1yLyPProLeuTyra1eSer1eA1yTPSerAaP1aProA1aSerTh1yA1yA 565
DB 1021 GCGGCAAGCTTGAAGATCTCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 566 SerVal1Th1eVal1AaenA1eA1yA1yA1yA1yA1yA1yA1yA1yA1yA1yA1y 585
DB 1081 TCCGTACCTTCTATACATGATCTGACCTTGTATTAACGCTTCAATGACACATAT 1140
QY 586 Val1yAaenA1yPheSerA1aProPheAaenA1yA1yA1yA1yA1yA1yA1yA1yA1y 605
DB 1141 GTTGAATATGATCTTATTCGCAATACATTAATTAATGAGGCGCAATTAACGTAAGAA 1200
QY 606 AenVal1Phe11eAenSerProG1nSerG1yThrTyra11eG1yVal1G1nA1yTyraPn 625
DB 1201 AATGATTTAATTAATGCGCACAAAGCGGAGATTAATGATGATGATGATGATGATGAT 1260
QY 626 ValProValG1yProG1nAaPheSer1eA1yA1yA1yA1yA1yA1yA1yA1yA1y 639
DB 1261 GTACCGGTTGACCAAGACCTTCTGTTGGCAATGTGAAC 1302

RESULT 14
US-10-385-662-1
/ Sequence 1, Application US/10385662
/ Publication No. US20040002432A1
/ GENERAL INFORMATION:
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SATO, TSUYOSHI
/ APPLICANT: SATO, KAZUHIRO
/ APPLICANT: SUMITOMO, NOBUYUKI
/ APPLICANT: IZAWA, YOSHIYUKI
/ APPLICANT: SAKETI, KATSUHIISA
/ APPLICANT: KOBAYASHI, TOHRU
/ APPLICANT: NOMURA, MASAFUMI
/ TITLE OF INVENTION: ALKALINE protease
/ FILE REFERENCE: 234938USO
/ CURRENT APPLICATION NUMBER: US/10/385,662
/ PRIOR FILING DATE: 2003-03-12
/ PRIOR APPLICATION NUMBER: JP 2002-081428
/ PRIOR FILING DATE: 2002-03-22
/ PRIOR APPLICATION NUMBER: JP 2002-165987
/ PRIOR FILING DATE: 2002-06-06
/ PRIOR APPLICATION NUMBER: JP 2002-304230
```

PRIOR FILING DATE: 2002-10-18
PRIOR APPLICATION NUMBER: JP 2002-304231
PRIOR FILING DATE: 2002-10-18
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1305
TYPE: DNA
ORGANISM: Bacillus sp. KSM-KP43
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1305)
OTHER INFORMATION:
US-10-385-662-1

Alignment Scores:

Pred. No.:	5,67e-189	Length:	1305
Score:	2183.00	Matches:	418
Percent Similarity:	99.3%	Conservative:	13
Best Local Similarity:	96.3%	Mismatches:	3
Query Match:	66.8%	Indels:	0
DB:	6	Gaps:	0

US-10-784-870-4 (1-639) x US-10-385-662-1 (1-1305)

QY 206 AaaspValaIaarglyIleValIaaspValaIaInserSerTyrglyLeuTy 225
DB 1 AATGAGTTGGCGGTGGAATTCAGAAAGCGAGTGGCTCAGAGCAGCTACGGGTTGTAT 60
QY 226 GlyInglyGlnIleValaIaValaIaaspThrglyLeuaspThrglyArgaspSer 245
DB 61 GGAACAAGGACAGATCTTACCGGTGCGATACAGGGCTGATACAGTGCATGACAGT 120
QY 246 SerMetHISgluaIaphaerpglyIleThraIaleuTyraIaleuIyargThraan 265
DB 121 TCGATGCAAGAACCTTCGCGGAAATTACTGCAATTAATGATGAGACGAGAAAT 180
QY 266 AaanaIaaspThraaspThraaspThraaspThraaspThraaspThraaspThra 285
DB 181 AATGCCAATGATACGAATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 286 AlaThrAsnIleGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSer 305
DB 241 TCCACCAATTAAGAAATGCGCTCAGGGAATCTAGTCTCAATCTATCATGATGATGAC 300
QY 306 SerGlyIleuGlyIleuPProSerAsnLeuGlnThleuPheSerGlnIaIaPheSer 325
DB 301 GGTGGGGGACTTGGAGGACTTCTTCAATCTGCAAACTTATTCAGCCAGCATACAGT 360
QY 326 AlaGlyIaIaargIleHISThrAsnSerTrpGlyAlaIaValaIaaspGlyValaTyrrThra 345
DB 361 GCTGGTGCAGAAATTCATCAAACTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 346 AaPserIa 365
DB 421 GATTTCAGAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 366 GlyaengIuaIa 385
DB 481 GGGATGAAAGACCGAAGCGGGAACCATCACTGCAAGGAGGAGGAGGAGGAGGAGGAG 540
QY 386 ThrValIa 405
DB 541 ACAATCGAGGCTACGGAAGAACTCCGCCCAAGCTTTGGGTCTTATGCGGACAAATACAC 600
QY 406 HISValaIa 425
DB 601 CATGGGACAGTTCTCTTCAAGTGAACGCAAAAGATGAGCGGATCAAAACCGGATGTC 660
QY 426 MetaIaPProGlyThraIleuSerAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 445
DB 661 ATGGACCGGGAAGCTTCACTACTACAGCAATCTTCTTTCACCGGATTCCTCTTC 720

QY 446 TrpAlaAsnHISaaspSerTyrrAlaTyrrMetGlyIleThraSerMetAlaThrProIle 465
DB 721 TGGGGAACCATGACGAAATATGATCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 466 ValaIa 485
DB 781 GTTGGTGGAAAGCTGGACAGCTTGTGAGCAATTTTGGAABAAACAGAGGATCACACA 840
QY 486 LysPProSerLeuLeuIleAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 505
DB 841 AAGCTTCTCTATTAAGCGGACGATGATGCGGCTGACGCTGACATCGGCTTGCTAC 900
QY 506 ProaengIuaengIlyTrpGlyArgValThleuaspIySerSerLeuaspValaIaTyrr 525
DB 901 CCGAAGGTTACCAAGATGGGAGAGATGACATGATTAATCCCTTAAGCTTGCCAT 960
QY 526 ValaengIuserSerAlaIa 545
DB 961 GTGAACGATCCAGTCTCTATTCACCAAGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
QY 546 AlaGlyIyPProLeuIleIleSerIleuValIlePserAlaIaIaIaIaIaIaIaIaIaIa 565
DB 1021 GCCGCAAGCCTTGAAATCTCCCTGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 566 SerValThleuValaIa 585
DB 1081 TCCGTAAGCTTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 586 ValGlyaaspPheSerAlaIa 605
DB 1141 GTAGGAATGATCTTACTTCCCATACAAATGATGATGATGATGATGATGATGATGATGAT 1200
QY 606 AaenValPheIleAsnSerProGlnSerGlyThraIleGlyValaIaIaIaIaIaIaIa 625
DB 1201 AATGATTTATTAATGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 626 ValProValIa 639
DB 1261 GTACCGTTGACCAAGACCTTCTGTTGCAATTTGANT 1302

RESULT 15

US-10-820-712A-21

Sequence 21, Application US/10820712A
Publication No. US2005026804A1

GENERAL INFORMATION:

APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Matsuyoshi
APPLICANT: Izawa, Yoshifumi
APPLICANT: Kobayashi, Tohru
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tetsuo
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-USO
CURRENT APPLICATION NUMBER: US/10/820, 712A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106708
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.2
SEQ ID NO 21

LENGTH: 1302
TYPE: DNA
ORGANISM: Bacillus NCIB1289
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1302)
OTHER INFORMATION: Gene name Al, protease Al

US-10-820-712A-21
Alignment Scores:
Pred. No.: 2,08e-183
Score: 2122.00
Percent Similarity: 97.7%

Length: 1302
Matches: 400
Conservative: 24

Best Local Similarity: 92.2% Mismatches: 10
 Query Match: 65.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-784-870-4 (1-639) x US-10-820-712A-21 (1-1302)

QY 206 AsnAspValAlaArgIleValIleValAlaAspValAlaGlnSerSerYrGlyLeuYr 225
 DB 1 AACGATGTTCCAGAGGCAATTGAAAGCCGATGTTCCAGAGCACTATGTTTAT 60
 QY 226 GlyGlyGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSer 245
 DB 61 GGGCAAGGGCAAGTGGTTCAGTAGCGGATACCGGATCGATACAGGGCTTAATGACGC 120
 QY 246 SerMetHisGlnAlaPheArgGlyValIleThrAlaLeuYrAlaLeuGlyValGthrAsn 265
 DB 121 TCATCATGAAACGGTTCAGAGAAAGATTACCGCAATATATGCCCTTGGAAACAAAC 180
 QY 266 AsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerValIleGlyAsnGly 285
 DB 181 AACGGCAATGATCAAAACGACACGGGACGATCTGCCGATCTGTTTAGAAACGAT 240
 QY 286 AlaThrAsnIleGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSer 305
 DB 241 ACTTCGAATTAAGGATGGCTCCACAAGCTTAATCTTCAATCTGTTATGACAGC 300
 QY 306 SerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrIleuPheSerGlnAlaPheSer 325
 DB 301 AATGGCGGACTTGGCGGACTGCTTCMAATGTAAGTACATTAATTCAGCCAGGCAATAGT 360
 QY 326 AlaGlyAlaArgIleHisThrAsnSerTrpGlyAlaAlaValAlaGlnIleAlaYrThrThr 345
 DB 361 GCCGGTCCAGATCCATACGAACTCATGGGAGCGCTTAACGAGCGTACACTACT 420
 QY 346 AspSerArgAsnValAspAspYrValArgIleAsnAspMetThrIleuPheAlaAla 365
 DB 421 GATTCAGAAACGTAGACGATTATGTTGTAATAATGATGGCGGTTCTTTTGGACGG 480
 QY 366 GlyAsnGlyLeuArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaValAsnAlaIle 385
 DB 481 GGTAACGAAGGGCGGATGGCGGACAAATCAATGCTCTGTAACCGGAATGCTATC 540
 QY 386 ThrValGlyAlaThrGlnAsnLeuArgProSerPheGlySerYrAlaAspAsnIleAsn 405
 DB 541 ACAGTAGGGGCAACAAACAACTGCCGCCAAGCTTGGATCTTAATGCTACAACTAAT 600
 QY 406 HisValAlaGlnPheSerSerArgIleProThrIleAspGlyArgIleuProAspVal 425
 DB 601 CATGTAGCAAGTTTCTCCCGGAGCTCAAAAGATGAGCTATCAAAACCGGACGT 660
 QY 426 MetAlaProGlyThrYrIleLeuSerAlaArgSerSerLeuAlaProAspSerPhe 445
 DB 661 ATGGCACCGAACAATTATTTATCGGCAAGATCTTCTTGGCTCCGACCTCTCATTC 720
 QY 446 TrpAlaAsnHisAspSerLeuYrAlaYrMetGlyGlyThrSerMetAlaThrProIle 465
 DB 721 TGGGCAACCATACAGCAAAATATGCTTAATGGGAGAACATCCATGGGACACCGATT 780
 QY 466 ValAlaGlyAsnValAlaGlnLeuArgGlyHisPheValIleAsnArgGlyIleThrPro 485
 DB 781 GTAAGCTGTACGTTGACAGTTACGTGAACATTTCTATCAAAAACAGAGGAATCACTCT 840
 QY 486 LysProSerLeuLeuValAlaAlaLeuIleAlaGlyAlaAlaAspValGlyLeuGlyYr 505
 DB 841 AAACCATCTTGTCTGAAGCAGCTTATATGCGGAGCAACTGATATCGGTCTTGGCTAT 900
 QY 506 ProAsnGlyAsnGlnIleTrpGlyArgValIleThrLeuAspLeuSerLeuAsnValAlaYr 525
 DB 901 CCGAGTGAACCAAGAGATGGGAAAGATGACATGACAAAGTCACTTAATGTAGCTTTC 960
 QY 526 ValAsnGlySerSerAlaLeuSerThrSerGlnIleValAlaThrYrThrPheThrAlaThr 545
 DB 961 GTAAATGAACAAAGCTTTTATCTACTAACCAAAAGGCTACGATTCAATTTACTGCAAA 1020

QY 546 AlaGlyLysProLeuLysIleSerLeuValIleTrpSerAspAlaProAlaSerThrThrAla 565
 DB 1021 TCAGGCAACCTTGAAGATTTCATTGGTTGGTCTGATGACCGGCAAGTACTTCCGCA 1080
 QY 566 SerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyThrArgYr 585
 DB 1081 TCGGTTACATGGTGAATGATCTGATCTGGTATTAACGCTCAAAATGGAACAAAGTAT 1140
 QY 586 ValGlyAsnAspPheSerAlaProPheAspAsnIleTrpAspGlyArgAsnValGln 605
 DB 1141 GTTGAACAGCACTTACTGCTCCCTATGATTAATTAACGGGATGACGTAACAAATGAGAG 1200
 QY 606 AsnValPheIleAsnSerProGlnSerGlyThrYrThrIleGlyValGlnAlaYrAsn 625
 DB 1201 AACGTGTTATCAATGCTCCGCAAAACGGAAACGTATACGTTGAGGTACAGGCTTACAAAT 1260
 QY 626 ValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
 DB 1261 GTTCCAAAGGGCGGCGGCTTTCTTGGCTATTTGTGAC 1302

Search completed: April 8, 2006, 04:36:19
 Job time : 1356 secs

This Page Blank (usp10)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 8, 2006, 04:18:10 ; Search time 2708 Seconds

(without alignments)
944.001 Million cell updates/sec

Title: US-10-784-870-4

Perfect score: 3267

Sequence: 1 MRKKVFLSLTAAALSTV.....EVQAYNPVGPQNFSLAIV 639

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ .p2n.model -DEV=xlp
-Q=/abse/ABSSWB.spool/US10784870/runat_07042006_143430_27776/app_query.fasta.1
-DB=Published Applications NA New -OPMT=fastcap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPTC=0 -LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62
-TRANS=human0.csl -LIST=45 -DCCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abss05p
-USER=US10784870.qden 1.1.987 @runat_07042006_143430_27776 -NCPU=6 -ICPU=3
-NO MAMP -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV TIMEOUT=120
-NARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA New:*
1: /SIDS5/ptodata/1/pubpna/US08 NEW PUB.seq:*
2: /SIDS5/ptodata/1/pubpna/US06 NEW PUB.seq:*
3: /SIDS5/ptodata/1/pubpna/US07 NEW PUB.seq:*
4: /SIDS5/ptodata/1/pubpna/PCT_NEW PUB.seq:*
5: /SIDS5/ptodata/1/pubpna/US09 NEW PUB.seq:*
6: /SIDS5/ptodata/1/pubpna/US10 NEW PUB.seq:*
7: /SIDS5/ptodata/1/pubpna/US10 NEW PUB.seq:*
8: /SIDS5/ptodata/1/pubpna/US10 NEW PUB.seq:*
9: /SIDS5/ptodata/1/pubpna/US10 NEW PUB.seq:*
10: /SIDS5/ptodata/1/pubpna/US10 NEW PUB.seq:*
11: /SIDS5/ptodata/1/pubpna/US11 NEW PUB.seq:*
12: /SIDS5/ptodata/1/pubpna/US11 NEW PUB.seq:*
13: /SIDS5/ptodata/1/pubpna/US11 NEW PUB.seq:*
14: /SIDS5/ptodata/1/pubpna/US11 NEW PUB.seq:*
15: /SIDS5/ptodata/1/pubpna/US60_NEW PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	315	9.6	3625	8 US-10-510-386-27	Sequence 27, Appl
2	315	9.6	4041	8 US-10-510-386-199	Sequence 199, App

3	292.5	9.0	3405	8 US-10-510-386-1	Sequence 1, Appl
4	260.5	8.0	1140	14 US-11-156-062-22	Sequence 22, Appl
5	245	7.5	1495	14 US-11-020-602-1	Sequence 1, Appl
6	244	7.5	4216	14 US-11-136-527-605	Sequence 605, Appl
7	243	7.4	4338	14 US-11-122-329-113	Sequence 113, App
8	242.5	7.4	1180	11 US-11-047-380-4	Sequence 4, Appl
9	242	7.4	4199	14 US-11-128-061-679	Sequence 679, App
10	242	7.4	4199	14 US-11-128-049-679	Sequence 679, App
11	238.5	7.3	5296	8 US-10-510-386-217	Sequence 217, App
12	226.5	6.9	824	14 US-11-156-062-13	Sequence 13, Appl
13	226.5	6.9	825	14 US-11-065-943-59	Sequence 59, Appl
14	225.5	6.9	824	14 US-11-156-062-11	Sequence 11, Appl
15	223	6.8	2760	13 US-11-008-331-1	Sequence 1, Appl
16	222.5	6.8	824	14 US-11-156-062-7	Sequence 7, Appl
17	221.5	6.8	824	14 US-11-156-062-3	Sequence 3, Appl
18	220.5	6.7	822	14 US-11-156-062-9	Sequence 9, Appl
19	219.5	6.7	824	14 US-11-156-062-5	Sequence 5, Appl
20	217	6.6	1437	9 US-10-932-182A-2353	Sequence 2353, Ap
21	217	6.6	1437	9 US-10-932-182A-2353	Sequence 2353, Ap
22	216.5	6.6	1437	9 US-10-932-182A-2353	Sequence 2353, Ap
23	216.5	6.6	824	14 US-11-156-062-15	Sequence 15, Appl
24	215.5	6.6	824	14 US-11-156-062-17	Sequence 17, Appl
25	211.5	6.5	824	14 US-11-156-062-1	Sequence 1, Appl
26	205	6.3	2772	8 US-10-510-386-21	Sequence 643, App
27	197.5	6.0	1476	9 US-10-932-182A-75743	Sequence 21, Appl
28	197.5	6.0	1476	9 US-10-932-182A-75743	Sequence 75743, A
29	197.5	6.0	2566	11 US-11-096-568A-24711	Sequence 24711, A
30	190	5.8	1908	9 US-10-932-182A-76207	Sequence 76207, A
31	190	5.8	1908	9 US-10-932-182A-76207	Sequence 76207, A
32	189	5.8	1815	9 US-10-932-182A-669	Sequence 669, App
33	189	5.8	1815	9 US-10-932-182A-669	Sequence 669, App
34	182	5.6	4856	14 US-11-136-527-3850	Sequence 3850, App
35	175.5	5.4	2251	14 US-11-128-061-10	Sequence 10, Appl
36	175.5	5.4	2251	14 US-11-128-061-10	Sequence 10, Appl
37	175.5	5.4	2388	14 US-11-218-986-1	Sequence 1, Appl
38	175.5	5.4	2543	11 US-11-096-568A-11250	Sequence 31250, A
39	175.5	5.4	4180	8 US-10-821-234-324	Sequence 324, App
40	171.5	5.2	4626	14 US-11-000-688-445	Sequence 445, App
41	168	5.1	2611	11 US-11-096-568A-11860	Sequence 31860, A
42	166.5	5.1	2527	14 US-11-128-061-1086	Sequence 1086, Ap
43	166.5	5.1	2527	14 US-11-128-061-1086	Sequence 1086, Ap
44	161	4.9	2706	11 US-11-096-568A-17893	Sequence 17893, A
45	158.5	4.9	2537	11 US-11-096-568A-23981	Sequence 23981, A

ALIGNMENTS

RESULT 1
US-10-510-386-27
Sequence 27, Application US/10510386
Publication No. US20050244922A1
GENERAL INFORMATION:
APPLICANT: Andersen, Jens Tonne
APPLICANT: Clausen, Ib Groth
APPLICANT: Jorgensen, Steen Troels
APPLICANT: Olsen, Peter Bjarke
APPLICANT: Rasmussen, Michael Dolberg
TITLE OF INVENTION: Improved Bacillus Host Cell
FILE REFERENCE: 10294.204-US
CURRENT APPLICATION NUMBER: US/10/510,386
CURRENT FILING DATE: 2004-10-04
NUMBER OF SEQ ID NOS: 248
SOFTWARE: PatentIn version 3.3
SEQ ID NO 27
LENGTH: 3625
TYPE: DNA
ORGANISM: Bacillus licheniformis
FEATURE:
NAME/KEY: CDS
LOCATION: (501)..(3122)
US-10-510-386-27
Alignment Scores:

Pred. No.: 9,06e-22 Length: 3625
 Score: 315.00 Matches: 161
 Percent Similarity: 41.4% Conservative: 92
 Best Local Similarity: 26.4% Mismatches: 234
 Query Match: 9.6% Indels: 125
 DB: 8 Gaps: 26

US-10-784-870-4 (1-639) x US-10-510-386-27 (1-3625)

```

QY 2 ArglyelyleValPheleuserValleu-SerAlaAlaIleleuserThrValAl 21
DB 1106 AGGAAAGAGCTGCTTC---CCTTTACTATAAGAGCTCATTCATCAGCGCAAAAT 1162
QY 21 AleuAnaAnProSerAlaGlyAspAlaArgThrPheAspLeuAsp----- 36
DB 1163 GCTCTTTAAACAATTCATAGAGACAGGTCTACAAAGACTGTGCAGCTGAAGCGCT 1222
QY 37 -PheylsGlyIleGlnThrThrAspValSerGlyPheSer---LysGlnArgGlnTh 55
DB 1223 GTTTGAGATGTGCGTAAAGAGAGACAGTGAACGCGATCTCATTAACAAATGATGATCA 1282
QY 55 rGlyAlaAlaPheleuValGlyuserGluAsnValLysLeuleuylsGlyLeuleu 75
DB 1283 AAAAGGATCAGTCCGCTGACAAAAGCTCCGCGCTCCGAGCGCTTGAAGAAA 1342
QY 75 slyVleuGluThrValProAlaAsnAsnLysleuHsIleValGlnPheAsnGly----- 93
DB 1343 ACAGCTGATCAAGTCCGAGAAAGC-----ATCGCATGAAACAATTAAACGCGCA 1396
QY 94 -----ProIleuGlnGluThrLysGlnLysleuGluThrThrGlyAlaLysIlele 111
DB 1397 AGTATCGGCTGCTGTAAGAAAGCGGGAGTGGGACAGCTTCAAGGCG----- 1448
QY 111 uAspTrIleProAspTrAlaTrIleValGluTrGlu-----GlyAs 126
DB 1449 -----CCGAAACCGTTACATCTAAATTCGAAAGAAAGGCAAAACCGGGATC 1498
QY 126 rValGlnSerLysValArgSerIleGluHsValGlnSerValGluPro----- 142
DB 1499 TTTCAATTTAAAGCCCAATCA---TCCGCGCTCCAGGCAATTAGACCCCTCGTAAAG 1555
QY 143 -----TyrLeuProLysTrIleAspProGlnLeuPheTh 155
DB 1556 CAAACCGCATTTAAAGATATGATGAGTGTGGAAATGAAGAAAGCGTTCTCCGAGT 1615
QY 155 rlyseGlylAserThrLeuValLysAlaLeuAlaLeuAspThrLysGlnAsnLysG 175
DB 1616 CAAACGCGCGGCAAGCAATATCAGCGCGCAGCC-----TCCAGATCGCCAGAT 1666
QY 175 uValGlnLeuArgGlyIleGlnGluIleAlaGlnTrValAla---SerAsnAspValH 194
DB 1667 GCGTAGAGGAAATTCGTGAAACAGGTTCAAGCAATTAAGACCTGTCAAGACACCA 1726
QY 194 sTrIleThrAlaLysProGluTrLysValLysAsnAspValAlaArgGlyIleVal 214
DB 1727 AAT-----CCATATCAATGCTCGCTCAAAATAACGGCAAAACCGTGGCGGA 1777
QY 214 sAlaAspValAlaGlnSerSerTrGlyLeuTrGlyGlnGlyGln----- 229
DB 1778 TGCTACATACAAATTGAACAGCTTCAGAACTGTGAAGGCAAAAGCTGAAGATAC 1837
QY 230 -IleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHsG 249
DB 1838 AGTATCCGCGCTCGTTGACACAGGCTTATCATACCTTGGCGATTAAAG----- 1889
QY 249 uAlaPheArgGlyLysIle-----ThrAlaLeuTrAlaLeuGlyArgThrAsnAs 266
DB 1890 -----GGCAGCGTCAAAAGAAAGCAAGCTATATACTATGCGCGGCGAGCGGA 1939
QY 266 nAlaAsnAspThrArgAsnGlyHsIleGlyThrHsValAlaGlySerValleuGly----- 283
DB 1940 TCGGATGATGACAAATGCGCAGGACACAGCTGTCAAGCATTTGACGCGCGCAAGA 1999
  
```

```

QY 284 -AsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMe 303
DB 2000 CAACCATTTTGTGATGCGCGGAAATCAATGTTATGCAAAATTCGCTGTCAAGGTCT 2059
QY 303 rAspSerSerGlyLysLeuGlyLys-LeuProSerAsnLeuGlnThrLeuPheSerGln 323
DB 2060 GGATTTCTTAGGC-----AGTGGAGATTCGAAACAGATTCGAAACCGCATATCTAT-G 2112
QY 323 lAsPheSerAlaGlyAlaArgIleHsIleThrAsnSerTrGlyAlaAlaValAsnGlyAla 343
DB 2113 CGCGGACACACGCTGCAAAAGTCATCAATTTAAAGCTTGGC-----GGCGCAT 2160
QY 343 yTrThrThrAspSerArgAsnValAspAspTrValArgGlyAsnAspMetThrIleLeuP 363
DB 2161 ACAGCGGCGTGAATGAAATATGCGCTTAAATATGCGGCTTAAATATGACAGATCGTGG 2220
QY 363 hEAlaAlaGlyAsnGluArgProAsnGlyLysIleSerAlaProGlyThrAlaLysVal 383
DB 2221 CGCGCACCGGAATGAC-----GAGATATGGAAGATTTCTACCTGATCTTGAAAT 2274
QY 383 sAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTrAlaAspAl 403
DB 2275 ATACGCTTTCAGTCCGGCGCACCAATATATCTC----- 2306
QY 403 sAlaAsnHsIleValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIleLysP 423
DB 2307 -----GATCTGTGTCGACTCATCTCAATATGAAAGTCTC----- 2345
QY 423 rAspValMetAlaProGlyThrTrIleLeuSerAlaArgSerSerLeuAlaProAsP 443
DB 2346 --GATATGATGCGCGCGGAAACCATATTT-----CMAAGCTCGTCCGACG 2391
QY 443 eSerPheThrAlaAsnHsIleAspSerLysTrAlaTrMetGlyLysIleSerMetAlaT 463
DB 2392 GGAAT-----GTCACTTATATAGAGGAAACATGATGCGG 2427
QY 463 hrProIleValAlaGlyAsnValAlaGlnLeuArgGlnHsIlePheValLysAsnArgGly 483
DB 2428 CGCGCACGTGCGAGCTCGACAGAGACTTCTT-----TTGCACAGATC 2472
QY 483 lThrProLysProSerLeuLeuLysValAlaLeuIleAlaGlyAlaAlaAspValGlyL 503
DB 2473 CGTCTTGAACCAAGCAATTCGAGACCTATTTAGCCGACAGACAGACAGATGTGGAT 2532
QY 503 eu-----GlyTrProAsnGlyAsn----- 509
DB 2533 TTGAAGACAGATATATCAACCGGATTTGACTGATATAGAACCGGCTGCACAA 2592
QY 510 -----GlnGlyTrGlyArgValAlaThrLeuAspLysSerLeuAsn 523
DB 2593 TTCCGAGTATGACTTCTCCGAGTGGAGAGGCTGAATGTTTTCATGACAGCGG 2652
QY 523 AlaIaTrValAsnGlnSerSerAlaLeuSerThrSerGlnLysAlaIleThrPheT 543
DB 2653 TTTTGAAGCTGAACATGAAGTTCATCCGTTTAAACCGGCATACGCGCATGACAGCA 2712
QY 543 hrAlaThrAlaGlyLysPProLeuLysIle 552
DB 2713 CAGCCAAAGCGGTGTGACGCTCAAAATC 2741
  
```

RESULT 2

US-10-510-386-199
 ; Sequence 199, Application US/10510386
 ; Publication No. US20050244922A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Andersen, Jens Tonne
 ; APPLICANT: Clausen, Ib Groch
 ; APPLICANT: Jorgensen, Steen Troels
 ; APPLICANT: Olsen, Peter Bjarke
 ; APPLICANT: Raasmussen, Michael Dolberg
 ; TITLE OF INVENTION: Improved Bacillus Host Cell
 ; FILE REFERENCE: 10294.204-US
 ; CURRENT APPLICATION NUMBER: US/10/510.386

CURRENT FILING DATE: 2004-10-04
NUMBER OF SEQ ID NOS: 248
SOFTWARE: PatentIn version 3.3
SEQ ID NO 199
LENGTH: 4041
TYPE: DNA
ORGANISM: Bacillus licheniformis
FEATURE:
NAME/KEY: CDS
LOCATION: (501) .. (3641)
US-10-510-386-199

Alignment Scores:
Pred. No.: 1,03e-21 Length: 4041
Score: 315.00 Matches: 161
Percent Similarity: 41.4% Conservative: 92
Best Local Similarity: 26.4% Mismatches: 234
Query Match: 9.6% Indels: 125
Gaps: 26

US-10-784-870-4 (1-639) x US-10-510-386-199 (1-4041)

QY 2 ArgLyLeuValPheLeuSerValLeu-SerAlaAlaAlaLeuSerThrValAl 21
DB 1106 AGGAAAGAGCTGCTTC---CCTTACTATAAGAGCTCCATTACAGCGCAAAAT 1162
QY 21 AlaLeuAsnProSerAlaGlyAspAlaArgThrPheAspLeuAsp----- 36
DB 1163 GCTCTTAAACAAATTCAGTAGAGACAGTGTCTACAAAGACTGTGACGTGAAGCGCT 1222
QY 37 -PheYsgLYLeuThrThrThrAspValSerGlyPheSer---LysGlnArgGlnTh 55
DB 1223 GTTTCAGATGCTGCTAAAAACGACAGTAGAGCGCATCTCAATTACGATGATGATCA 1282
QY 55 rGlyAlaAlaAlaPheLeuValGluSerGluAsnValLysLeuLeuYsgLYLeuY 75
DB 1283 AAAAGGATCAGTCGCTGTACGAAACAGCTCGCGGTCCGAGCGTTGAAAGAA 1342
QY 75 GlyLeuGluThrValPheAlaAsnAsnLysLeuHisLeuAlaGlnPheAsnGly---- 93
DB 1343 ACAGCTGATCAGTGCAGCGGAAAGAC---ATCGCATCGAACATTAACAGGACGAA 1396
QY 94 -----ProLeuGluGluThrLysGlnLysGluThrThrGlyValAlaLysIle 111
DB 1397 AGTATCGCTGTCTTGAAAGACGGGATGCGGACGCTTCAACAGCGG----- 1448
QY 111 uAspYrIleProAspYrIleArgYrIleValGluYrGlu-----GlyAs 126
DB 1449 -----CCGAAACCGCTTACATCTGTAATAATGAAAGGACGAAACCGGGATC 1498
QY 126 pValGlnSerLysValArgSerIleGluHisValGluSerValGluPro----- 142
DB 1499 TTTCAATCTAAAGCCCAATCA---TCCGCGTCCAGGCGCATAGACCCCTCGTAAG 1555
QY 143 -----TyrLeuProLysYrIleAspProGlnLeuPheTh 155
DB 1556 CAAACCGCATTTAAAGATATGATGCTGTGGAAATGAAAGAAAGCGCTTCCGAGAT 1615
QY 155 rLyGlyAlaSerThrLeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysG 175
DB 1616 CAAAGCGGCGGCAAAACATATCAGCGCGGAGCC-----TCCAGATCCGCAAGAT 1666
QY 175 uValGlnLeuAspGlyrIleGluGluThrLeuAlaGlnYrValAla---SerAsnAspValH 194
DB 1667 GCCTGAAGTGAATTTGTCGACAGGTTGACGAAATGAAAGACACTGTCAAGACACCCA 1726
QY 194 sYrYrIleThrAlaLysProGluYrLysValMetAsnAspValAlaArgGlyrIleVal 214
DB 1727 ATATY-----CAATATCAATGTGCTCAAAATAACGCGAAACCGTGTGCGAA 1777
QY 214 sAlaAspValAlaGlnSerSerYrGlyLeuYrGlyGlnGlyGln----- 229
DB 1778 TGTGTGATATACATTTTGAACAGCTTCAAGAGCTGATGAAAGGAAAGAGTGAAGATAC 1837

QY 230 -IleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisG 249
DB 1838 AGTAATCGCGCTGTTGACACAGCGGTATATCATACCTTGGCGAATTTAAGC----- 1889
QY 249 uAlaPheArgLYrIle-----ThrAlaLeuYrAlaLeuGlyArgThrAsnAs 266
DB 1890 -----GCGAGCGTCAAAAAGACGAGCTTAATCTATGTGCGCGCACGCGCGGA 1939
QY 266 nAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGly----- 283
DB 1940 TGGGATGATGACAAATGCGCACGCGCACACGCTGACGAGCATATTCAGCCGCAAGA 1999
QY 284 -AsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIle 303
DB 2000 CAACCAATTTTCATGAGCGGGAATCAATGCTTATGCAAAATCTCTGTCAAAAGCTT 2059
QY 303 AspSerSerGlyGlyLeuGlyYrLysProSerAsnLeuGlnThrLeuPheSerGln 323
DB 2060 GATTTCTTACGCG-----AGTGAATACGGAACGATTTGCAACGCGCATCTAT-G 2112
QY 323 lAspSerAlaGlyAlaArgGlyLeuThrAsnSerTrpGlyAlaAlaValAsnGlyAla 343
DB 2113 CCGCGACACGCGTGCAGAAAGTCATCAATTAAGTCTTGCGC-----GGGCGAT 2160
QY 343 YrThrThrAspSerArgAsnValAspAspYrValArgLysAsnAspMetThrIleLeu 363
DB 2161 ACAGCGGCTGATGAAATATGCGCTTAAATATGCGGCTTCAAAATGAGACATGCTGT 2220
QY 363 hAlaAlaGlyAsnGluArgProAsnGlyYrThrIleSerAlaProGlyThrAlaLys 383
DB 2221 CCGCACCGGAATATGAC-----GGAGTATCGAGATTTCTTACCTGTCATCTTCGAAT 2274
QY 383 sAlaIleThrValGlyAlaArgGluAsnLeuArgProSerPheGlySerYrAlaAsp 403
DB 2275 ATACGCTTCAAGTCGCGGAGCACAATATATCTC----- 2306
QY 403 sNleAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIleLys 423
DB 2307 -----GATCTTGTCTGAGCTCACTCAATATGAAAGGTC----- 2345
QY 423 roAspValMetAlaProGlyThrYrIleLeuSerAlaArgSerLeuAlaProAsp 443
DB 2346 --GATATGCTGCGCGGAGACGATAT-----CCAGCCTTCGTTCCGAGC 2391
QY 443 eSerPheThrAlaAsnHisAspSerYrAlaYrMetGlyYrThrSerMetAla 463
DB 2392 GGAT-----GTCACTATATGAGCGGAACATGATGCGCG 2427
QY 463 h-ProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPheValLysAsnArgGly 483
DB 2428 CCGCGACCTGCGACGTCGACGACGACTTCT-----TTGTCAACAATC 2472
QY 483 lThrProLysProSerLeuLeuLysValAlaLeuAlaGlyAlaAlaAspValGly 503
DB 2473 CGTCTTGAACCAAAACAAATCGCAAGCTTATGACGAGACGACGAGATGTGCAT 2532
QY 503 eu-----GlyYrProAsnGlyAsn----- 509
DB 2533 TTGAAGACGAGATATATCAAAACCGGATTAATGACTGATATGAAACCGGTCACAAA 2592
QY 510 -----GlnGlyTrpGlyArgValThrLeuAspLysSerLeuAsn 523
DB 2593 TTCCCGGATATGACTTCTCGCGGTGGGGAAGCGTGAATGTTTTCATGACGACGAG 2652
QY 523 sAlaIleYrValAsnGluSerSerAlaLeuSerThrSerGlnYrValThrYrThrPhe 543
DB 2653 TTTTGAAGCTGAACATGAAGTTCAATCCCGTTTAAACCGCATACGCGAGTGAAGGCA 2712
QY 543 hAlaThrAlaGlyLysProLeuYrIle 552
DB 2713 CAGCCAAAGCGGTGACGTCAAATC 2741

```
RESULT 3
US-10-510-386-1
/ Sequence 1, Application US/10510386
/ Publication No. US20050244922A1
/ GENERAL INFORMATION:
/ APPLICANT: Andersen, Jens Tonne
/ APPLICANT: Andersen, Ib Groth
/ APPLICANT: Clausen, Steen Troels
/ APPLICANT: Jorgensen, Peter Bjarke
/ APPLICANT: Rasmussen, Michael Dolberg
/ TITLE OF INVENTION: Improved Bacillus Host Cell
/ FILE REFERENCE: 10294.204-US
/ CURRENT APPLICATION NUMBER: US/10/510,386
/ CURRENT FILING DATE: 2004-10-04
/ NUMBER OF SEQ ID NOS: 248
/ SOFTWARE: Patentin version 3.3
/ SEQ ID NO 1
/ LENGTH: 3405
/ TYPE: DNA
/ ORGANISM: Bacillus licheniformis
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (501)..(2906)
/ US-10-510-386-1

Alignment Scores: %
Pred. No.: 1,77e-19 Length: 3405
Score: 292.50 Matches: 166
Percent Similarity: 34.2% Conservative: 84
Best Local Similarity: 22.7% Mismatches: 219
Query Match: 9.0% Indels: 263
DB: 8 Gaps: 33

US-10-784-870-4 (1-639) x US-10-510-386-1 (1-3405)

QY 126 AppValGlnSerLeuValArgSerIleGluHsValGluSerValGluProTyrLeuPro 145
DB 645 GATGTGACGCTGTGATTAACAGACGACGCTTATCGTGAACTGAAAGAAAGCGCTTGGC 704
QY 146 LysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThrLeuValLysAlaLeu 165
DB 705 GAAGCAAAAGCCGACGAGAAAACAAACGAGGCTTCTTAAACAGCTTGAAACAA 764
QY 166 AlaLeuAspThrLysGlnAsn-----AsnLysGluValGln-----LeuArgGly 180
DB 765 GCGCTGAAACACTTAATAAAGCAAGTAACCCGGAATACGACCGTATTTCCGGC 824
QY 181 Ile-----GluGluIleAlaGlnTyrValAlaSerAsnAspValHis 194
DB 825 TTTTCTATGAAACTCCGCGCCAGTGAATTCCAAAGCTGCTCGCGTCAAGGAATTA 884
QY 195 TyrIle-----ThrAlaLysProGlu----- 201
DB 885 GCGGTTTATCCAAACGCAACTTACAAACCTGACACGCTTAAGAAAAGACGTCACATT 944
QY 202 -----TyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAsp 216
DB 945 GCAGCAGACGCCATTATCCACAGATG---GATTAAGATGCCCCCTTCATCGAGCGCAT 1001
QY 217 ValAlaGlnSerSerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThr 236
DB 1002 CAGCGATGAAATGCG---GGCTATACCGGTAAAGGCATTAAAGTACGCGTTATGATACG 1058
QY 237 GlyLeuAspThrGlyArgAsnAsp---SerSerMetHisGluAlaPheArgGlyLys--- 254
DB 1059 GCGCGTGAATTAACCCATCCGACCTTAAGAAATTAATTTCCGCCCATATTAAGGCGTANGAC 1118
QY 255 ---IleThrAlaLeuTyrAlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGly--- 272
DB 1119 TTGTGTGATATGATATGATATCAAGAAACACCGACAGGCAACCGCGCGGAGGAGCG 1178
QY 273 -----HisGlyThrHisValAlaGlySerValLeuGlyAlaGlyAlaThrAsnLysGly 290
DB 273 -----HisGlyThrHisValAlaGlySerValLeuGlyAlaGlyAlaThrAsnLysGly 290
```

```
DB 1179 ACTGATCAACGCGCATCATGTTGACAGAAACAATCGCCGGAACGACAGATT---AAAGGC 1235
QY 291 MetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerGlyGlyLeuGly 310
DB 1236 GTGGCGCGTGAAGCAACGCTTCTTGTCTTAACCGCGTCTCGGCCCGCGGCGC-----TCA 1289
QY 311 GlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAlaArgIle 330
DB 1230 GGCACAAACCGAAGAAATGATATGCGCGCATTTGAAAAGCCGTTGCGACAGCGAAGAAATG 1349
QY 331 HisThrAsnSerTyrGlyAlaAlaValAsnGlyAla---TyrThrThaAspSerArgAsn 349
DB 1350 ATGAACCTCTCATTTGGGAAATTCGTTGACAGCCCTGATTTATGCAACAAACATC----- 1403
QY 350 ValAspAspTyrValArgLysAsnAspMetThrIleLeuPheAlaAlaGlyAsnGluArg 369
DB 1404 GCACGTGACTGGCGCATGCGTGAAGCGGTGTGCGCGTTACGTCAACAGTAAACGCGGA 1463
QY 370 ProAsnGlyGlyThrIleSerAlaProGlyThrAlaLysAsnAlaIleThrValAla 389
DB 1464 CCGAAGAACTGACGCGTGGCTCGCCGCGGAACATCAAGGTTGCGATTTCCGTCGCGCA 1523
QY 390 ThrGlu-----AsnLeuArgProSerPheGlySer----- 399
DB 1524 TCACAGCTTCGCTAATAAGATATTCGCTGACGCTTCTTGTATTCATACGAAAGGTG 1583
QY 399 ----- 399
DB 1584 ATGGATATCAAGAGAAAGAAAGATCTTGAAAGCATTTGAACGCTCAAGAGTGAATCGTT 1643
QY 399 ----- 399
DB 1644 GAAGCCGCGCTCGACAAAGCTGACGATTTCTCAGGCAAAAGATGAAAGAAAGTTGCG 1703
QY 400 -----TyrAlaAspAsnIleAsnHis----- 406
DB 1704 GTTATTCAGAGAGCGTCATACATTTGTTGATTAAGCTGAAATTCGCAAAAAGCCGCGC 1763
QY 406 ----- 406
DB 1764 GCTATCGAGCGGTGATTTACATATATGCAACAGGGAATTCGAGCAAAAGTCATGCGG 1823
QY 406 ----- 406
DB 1824 ATGGCCGTACGACCGTCAAAATTTCTCAAAAGAGAGGCGCAAAAGCTTTTCAACAGATC 1883
QY 406 ----- 406
DB 1884 AAAGAAAGGAAACACTCCGTTGCTTTCTTCAATTTGCAAAAAGCTGGGGAAACA 1943
QY 407 ValAlaGlnPheSerSerArgGlyProThrLysAsp---GlyArgIleLysProAspVal 425
DB 1944 ATTCCCTCATCTTCGTCGCCGCGGCTGTCAATGATCATGATGATTAACCTGACGTT 2003
QY 426 MetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerPhe 445
DB 2004 TCAGCGCCAGCGCGTCAACATCTGTACGACCATTCGAC----- 2042
QY 446 TrpAlaAsnHisAspSerLys-----TyrAlaTyrMetGlyGlyThrSerMet 461
DB 2043 -----CACGATCGAAGAAACCGGTACGTTACGTTCAAAACAGGGAACAAGCATG 2093
QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPheValLysAsnArg 481
DB 2094 GCTTCCCGCATGTGCGGGAGCGGACGACATTAAGAG----- 2135
QY 482 GlyIleThrPro-----LysProSerLeuLeuLysAlaIleLeuIleAlaIle 499
DB 2136 ---GCCAAACCGGATTCGACGCTTGAGCAAAATCAAAAGCGTACTGATGAATACGGCGAA 2192
QY 500 AspVal-----GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgVal 515
DB 2193 AAACTCGCGATGAAGAAATGGAAGCCCTTCCTCAAAATGCGAAGAGAGCGGACGATC 2252
```

QY 516 ThrLeuApblySerLeuAenValAlaTyrValaAngluserSerAlaLeuSerThrSer 535
DB 2253 CGCATCATGAGGCCCTT-----AAAGCTTCATTCGATTTGTAACGCCG 2294
QY 536 GlnLysAlaThrTyr--ThrpheThrAlaThrAlaGlyLysProLeuLys----- 551
DB 2295 GGGAGCCATTCATACGGAACATTTCTGAAGACAAAGGAAACAGACGAAACAAAGCG 2354
QY 551 ----- 551
DB 2355 TTCACGATTGAAACCTTTCTTCACACGAAAGACCTATCAGCTCGAATACCTTTAA 2414
QY 552 -----1LeSerLeuValTyrSerAsp-----AlaProAlaSerThrThr-- 564
DB 2415 GGAACGGGCATTCACGGTATCAGAACGGAACGAGTCGTGTACCGGCATCAACAGCGT 2474
QY 565 -----AlaSerValThrLeuValaAenApblyAenValLleThrAlaProAngly 582
DB 2475 AAAGGAGCGGCAAAAGTAACCGTCATTTCCGCGAAACGAAAGCAGCAGCATTTGAAGGC 2534
QY 583 ThrArgTyrValGlyLysAenPheSerAlaProPheAenAenTyrAspGlyArgAsn 602
DB 2535 ACCGTTTACATCCGTGAA-----GACGGAAGAA 2564
QY 603 AenValGlu-----AenValPheLleAenSerPro----- 612
DB 2565 GTCCGCGAATCCCGCTCCTATGATGTCMAAGACGACATCACCGCGCTACATCC 2624
QY 613 -----GlnSerGlyThrTyrThrLleGluValGlnAlaTyrAsn 635
DB 2625 GTAACGTTGAACCGGAGACGAAAGCAGGCGCTTACGATCGAA-----GCCTAC-- 2675
QY 626 ValProValGlyProGlnAenPheSerLeuAlaIle 637
DB 2676 CTGCGCGGCGGCGCTGAAGAGCTCGCATTTCTGCTC 2711

RESULT 4
US-11-156-062-22
Sequence 22, Application US/1156062
Publication No. US20050281773A1
GENERAL INFORMATION:
APPLICANT: Wieland, Susanne
APPLICANT: Polanyi-Bald, Laura
APPLICANT: Prueßer, Inken
APPLICANT: Stehr, Regina
APPLICANT: Maurer, Karl-Heinz
TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
FILE REFERENCE: HEK-0134 / H5698
CURRENT APPLICATION NUMBER: US/11/156, 062
PRIOR FILING DATE: 2005-06-17
PRIOR APPLICATION NUMBER: PCT/EP2003/014127
PRIOR FILING DATE: 2003-12-20
PRIOR APPLICATION NUMBER: DE 102 60 903.9
PRIOR FILING DATE: 2002-12-20
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin version 3.3
SEQ ID NO 22
LENGTH: 1140
TYPE: DNA
ORGANISM: Bacillus licheniformis
FEATURE:
NAME/KEY: CDS (1) .. (1137)
LOCATION: (1) .. (1137)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (316) .. (1140)
US-11-156-062-22

Alignment Scores:
Pred. No.: 9.52e-17 Length: 1140
Score: 260.50 Matches: 120
Percent Similarity: 36.8% Conservative: 62

Best Local Similarity: 24.3% Mismatches: 157
Query Match: 8.0% Indels: 155
DB: 14 Gaps: 19

US-10-784-870-4 (1-639) x US-11-156-062-22 (1-1140)

QY 1 MetArgLysLysValPheLeuSerValLeuSerAlaAlaIleLeuSerThrVal 20
DB 4 ATGAGGAAAGAGTTTGGCTTGGATGACGCTTCATGCTCGTTCCAGATG 63
QY 21 AlaLeuAenAen-----ProSerAlaGlyAspAlaArgThrPheAenLeuAenPheLys 38
DB 64 GCATTCAGCATTCCTCCCTCTGCTGCTCAACCGCGGAAAGTTGAAAGATTAAT-- 120
QY 39 GlyLleGlnThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAla 58
DB 121 -----ATTGTCGATTT-----AAGTCAGAGTCGAA 147
QY 59 AlaPheLeuValGluSerGluAenValLysLeuLeuLysGlyLeuLeuLysLysLeu 78
DB 148 ACCGATCTGTCAAAAGACATCATCAAGAGCGCGGAAAGTGCACAG----- 201
QY 79 ThrValProAlaAenAenLysLeuHsileValGlnPheAenGlyProLleLeuGln 98
DB 202 -----CACTTTGA-----ATCATCAACGC 222
QY 99 ThrLysGlnLysLeuGlnThrThrGlyAlaLysLleLeuAenPyrLleProAspTyrAla 118
DB 223 GCAAAAGCGAGCTACAAAGAGAGCGCTTAAGGAATGCAAAATGATCCGATGTCCT 282
QY 119 TyrLleValGlnTyrGlnLysPheValGlnSerLysValArgSerLleGlnHsileVal 138
DB 283 TATGTGAAGAG-----GATCATGTG-- 303
QY 139 SerValGluProTyrLeuProLysTyrLysLleAspProGlnLeuPheThrLysGlyAla 158
DB 303 ----- 303
QY 159 SerThrLeuValLysAlaLeuAlaLeuAenPheThrLysGlnAenAenLysGluValGlnLeu 178
DB 303 ----- 303
QY 179 ArgGlyLleGlnLysLleAlaGlnTyrValAlaSerAenAenLysTyrLleThrAla 198
DB 304 -----GCCCATGCGCTACGGCA 321
QY 199 LysProGluTyrLysValMetAenAenValAlaArgGlyLleValLysAlaAenValAla 218
DB 322 ACCGTTCTTACGCGAT-----CTCTCATTAAGCGGACAAAGTG 363
QY 219 GlnSerSerTyrGlyLeuTyrGlyGlnGlyGlnLleValAlaValAlaAenPheThrGlyLeu 238
DB 364 CAGGCTCAA--GGCTTAAGGAGCGAATGTAAAGTACGCTCGATACAGAAATC 420
QY 239 AspThrGlyArgAenAenSerSerMetHsileValAlaPheArgGlyLysLleThrAlaLeu 258
DB 421 CAAGCTTCATCCGACCTTGAAACCTGACGCGCGA-----GCAGC 462
QY 259 TyrAlaLeuGlyArgThrAenAenAlaAenAenPheThrAenGlyHsileGlyThrHsileVal 278
DB 463 TTGTGGCTGGCGAAGCTTAAAC--ACCGACGGCAACGACACGACATGTTGCC 519
QY 279 GlySerVal-----LeuGlyAenGlyAlaThrAenLysGlyMetAlaProGlnAlaAen 296
DB 520 GGTACAGTACGTCGCGTGCATTAACAAGGGGTATTAAGGGGTGGCGGACGCGATAC 579
QY 297 LeuValPheGlnSerLleMetAenAenSerGlyGlyLysLeuGlyLysLeuProSerAenLeu 316
DB 580 TTGTACGCGGTTAAAGTACGATTCAGACGGA-----AGCGGATCATACAGCGCGCAT 633
QY 317 GlnThrLeuPheSerGlnAlaPheSerAlaGlyAlaArgLleHsileThrAenSerTyrGly 336
DB 634 GTAAGCGAATGAGTGGGCGACAAACGAGCATGATTTATGACCTTGAG 693

```

Qy 337 AlaAlaValAsnGlyValAlaThrThrThrAspSerArgAsnValAspAspTyrValAlaArgLys 356
    |||
Db 694 GGAGCA-----TCAGGCTGACAGCGATGAAACAGGACGTCAATGCAATGATGCA 744
Qy 357 AsnAspMetThrIleLeuPheAlaAlaGlyAsnGlyArgProAsnGly-----GlyThr 374
    |||
Db 745 AGAGGGGTGTGCTGTAGCTGCAGCAGGGAACAGCGGATCTTCAGAGAAACGATATCA 804
Qy 375 IleSerAlaProGlyThrAlaValAsnAlaIleThrValGlyAlaThrGluAsnLeuArg 394
    |||
Db 805 ATAGGCTATCTCGCGAAATACGATTTGTCATCGCTGTGGTGGCG----- 849
Qy 395 ProSerPheGlySerTyrAlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGly 414
    |||
Db 850 -----GTAGACTTAAACAGACAGACAGCTTCATTTTCAGCGTGGCA 891
Qy 415 ProThrLysAspGlyArgIleLysProAspValMetAlaProGlyThrTyrIleLeuSer 434
    |||
Db 892 -----GCAGAGCTTGAAGTCATGCTCTCGGCGCAGCGGATATACGC 933
Qy 435 AlaArgSerSerLeuAlaProAspSerSerPheThrAlaAsnHisAspSerLysTyrAla 454
    |||
Db 934 ACTTACCCACG-----AACACTTATGCA 957
Qy 455 TyrMetGlyGlyThrSerMetAlaThrProIleValAlaGlyAsnValAla----- 471
    |||
Db 958 ACATTGACGAAACGTCATGCTCTCTCTCATGTAGCGGAGCAGCAGCTTGTGATCTTG 1017
Qy 472 -----GlnLeuArgGluHisPheValLysAsnArg 481
Db 1018 TCMAAACATCCGACCTTCAGCTTCACAAATCCGGAACCGT 1059

```

RESULT 5
US-11-020-602-1

Sequence 1, Application US/11020602
Publication No. US20060024764A1

GENERAL INFORMATION:

APPLICANT: Batell, David

APPLICANT: Harding, Fiona

TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND

TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME

FILE REFERENCE: GCS27C2

CURRENT APPLICATION NUMBER: US/11/020,602

PRIOR FILING DATE: 2004-12-22

PRIOR APPLICATION NUMBER: US 09/500,135

PRIOR FILING DATE: 2000-02-08

PRIOR APPLICATION NUMBER: US 09/060,872

NUMBER OF SEQ ID NOS: 240

SOFTWARE: Patent Ver. 2.1

SEQ ID NO 1

LENGTH: 1495

TYPE: DNA

ORGANISM: Bacillus amyloliquefaciens

FEATURE:

NAME/KEY: mat peptide

LOCATION: (417)..(1495)

FEATURE:

NAME/KEY: CDS

LOCATION: (96)..(1244)

FEATURE:

NAME/KEY: misc feature

LOCATION: (582)..(584)

OTHER INFORMATION: The num at positions 582 through 584 which in a

OTHER INFORMATION: preferred embodiment (aac) is to code for

OTHER INFORMATION: asparagine, but which may also code for proline.

```

FEATURE:
NAME/KEY: misc feature
LOCATION: (597)..(599)
OTHER INFORMATION: The num at positions 597 to 599 which in a
OTHER INFORMATION: preferred embodiment (aac) is to code for
OTHER INFORMATION: asparagine, but which may also code for aspartic acid.
NAME/KEY: misc feature
LOCATION: (678)..(680)
OTHER INFORMATION: The num at positions 678 through 680 which in a
OTHER INFORMATION: preferred embodiment (gca) is to code for
OTHER INFORMATION: alanine, but which may also code for serine.
NAME/KEY: misc feature
LOCATION: (681)..(683)
OTHER INFORMATION: The num at positions 681 through 683 which in a
OTHER INFORMATION: preferred embodiment (tca) is to code for serine,
OTHER INFORMATION: but which may also code for alanine.
NAME/KEY: misc feature
LOCATION: (708)..(710)
OTHER INFORMATION: The num at positions 708 through 710 which in a
OTHER INFORMATION: preferred embodiment (gct) is to code for
OTHER INFORMATION: alanine, but which may also code for aspartic acid.
NAME/KEY: misc feature
LOCATION: (711)..(713)
OTHER INFORMATION: The num at positions 711 through 713 which in a
OTHER INFORMATION: preferred embodiment (gac) is to code for
OTHER INFORMATION: aspartic acid, but which may also code for alanine.
NAME/KEY: misc feature
LOCATION: (888)..(890)
OTHER INFORMATION: The num at positions 888 through 890 which in a
OTHER INFORMATION: preferred embodiment (acc) is to code for
OTHER INFORMATION: threonine, but which may also code for serine.
NAME/KEY: misc feature
LOCATION: (891)..(893)
OTHER INFORMATION: The num at positions 891 through 893 which in a
OTHER INFORMATION: preferred embodiment (tcc) is to code for
OTHER INFORMATION: serine, but which may also code for threonine.
NAME/KEY: misc feature
LOCATION: (1167)..(1169)
OTHER INFORMATION: The num at positions 1167 through 1169 which in
OTHER INFORMATION: a preferred embodiment (gaa) is to code for
OTHER INFORMATION: glutamic acid, but which may also code for glutamine.
US-11-020-602-1

```

Alignment Scores:

Pred. No.: 5.27e-15

Score: 245.00

Percent Similarity: 34.2%

Best Local Similarity: 23.5%

Query Match: 7.5%

DB: 14

Gaps: 18

US-10-784-870-4 (1-639) x US-11-020-602-1 (1-1495)

```

Qy 1 MetArgLysLysValPheLeuSerValLeuSerAlaAlaIleLeuSerThrVal 20
    |||
Db 96 ATGAGAGCAAAAGATGATGATGCTGCTTTCCTTACGCTTAACTTACATG 155
Qy 21 AlaLeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIle 40
    |||
Db 156 GCGTTCGACAGCATCTCTGCCAGGCG----- 185
Qy 41 GlnThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaPhe 60
    |||
Db 186 -----GCAGGGAATCAACGGGAAAG----- 209
Qy 61 LeuValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLysLeuGluThrVal 80

```

```

Db      210 -----AAATATATGTCGGGTTTAAACAGCAATGAGCAGATG 248
Qy      81 ProAlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlnLutThrLys 100
Db      249 AGCGCCGCTAAGAGAAAGATGTCATTCTGAAAAGGC-----287
Qy      101 GlnLysLeuGlnLutThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIle 120
Db      287 -----287
Qy      121 ValGlnTyrGlnLysAspValGlnSerLysValArgSerIleGlnHisValGlnSerVal 140
Db      288 -----GGAAAGTGCAGAAAGCAATCAATATATGATGAC-----320
Qy      141 GluProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThr 160
Db      321 -----GCAGCTTCAGCT 332
Qy      161 LeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGlnValGlnLeuAspGly 180
Db      333 ACATTAAACGAAAAGCTGTA-----AAAGAAATGAAAAGACCCGAGCGCTTAC 386
Qy      181 IleGlnGlnLysIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaLysPro 200
Db      387 GTTGAAGAA-----GATCAGTAGCA-----CATCGTAGCGGAGCTCCGTG 428
Qy      201 GluTyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSer 220
Db      429 CCTTACGGCGTATCACAA-----ATTAAGCCCT--GCTTCGAC 467
Qy      221 SerTyrGlyLeuTyrGlyGlnGlyIleValAlaValAlaAspThrGlyLeuAspThr 240
Db      468 TCTCAAGCTTACACCTGATCAAAATGTTAAAGCGGTATTCACAGCGGTATCGAT--524
Qy      241 GlyArgAsnAspSerSerMetHisGlnAlaPheArgGlyLysIleThrAlaLysTyrAla 260
Db      525 -----TCTCTCATCTCTGATTAAAGTAGACAGCGGAGCCAGACATGCTT 569
Qy      261 LeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySer 280
Db      570 CCTTCTGAACANNNNNTTCCAAAGACNNNNACTCTCAGGAACTCAGCTTGGCGGCACA 629
Qy      281 Val-----LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuVal 298
Db      630 GTTCCGCGCTTATATACCTCAATCGGTATTAAGCGGTGCGCAAGCANNNNCTTTAC 669
Qy      299 PheGlnSerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThr 318
Db      690 GCTGTAAAGTTCTCGTNNNNNGCT-----716
Qy      319 LeuPheSerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyrGlyAlaAla 338
Db      717 -----TCCGGCAATACAGCTGATC 737
Qy      339 ValAsnGlyAlaTyrThrAspSerArgAsnValAsp-----351
Db      738 ATTAAACGAATCGAGTGGCGGATCGCAACAAATATGACGTTATTAACATGACCTCGGC 797
Qy      352 -----AspTyrValArgLysAsnAspMet 359
Db      798 GACACTTTCGTTCTGCTTAAAGCGGAGCTTAAAGCCGCTGATCCGCGCTC 857
Qy      360 ThrIleLeuPheAlaAlaGlyAsnGlyArgProAsnGly-----GlyThrIleSerAla 377
Db      858 GTATGCGGTGGCGGACCGGTAAAGAGCANNNNNGGAGCTCAAGCAGACAGTGGCTAC 917
Qy      378 ProGlyThrAlaLysAsnAlaIleThrValGlyAlaThrGlnAsnLeuArgProSerPhe 397
Db      918 CCGGTAAATACCTTCTGTCATGAGTAGAGCGCT-----953
Qy      398 GlySerTyrAlaAspAsnIleAsnHisValAlaGlnPheSerSerArgLysProThrLys 417

```

```

Db      954 -----GTTGACAGCAGCAACAAAGACATCTTTCTCAAGCGTAGACCT-----998
Qy      418 AspGlyArgIleLysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSer 437
Db      999 -----GAGCTTATGATGATGACGACCTGGC-----GTATCTATCAAAAGC 1037
Qy      438 SerLeuAlaProAspSerSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMetGly 457
Db      1038 ACCCTTCT-----GAAACAAATACGGGGGCTACAC 1070
Qy      458 GlyThrSerMetAlaThrProIleValAlaGlyAsnValAla 471
Db      1071 GGTACGTCAATGCAATCTCCGACGTTGCCGAGCGGCTGCT 1112

RESULT 6
US-11-136-527-605
; Sequence 605, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounth, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIORITY FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIORITY FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 605
; LENGTH: 4216
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-605

Alignment Scores:
Pred. No.: 2,35e-14 Length: 4216
Score: 244.00 Matches: 138
Percent Similarity: 38.3% Conservative: 87
Best Local Similarity: 23.5% Mismatches: 233
Query Match: 7.5% Indels: 130
DB: 14 Gaps: 27

US-10-784-870-4 (1-639) x US-11-136-527-605 (1-4216)
Qy      87 HisIleValGlnPheAsnGlyProIleLeuGlnLutThrLysGlnLysLeuGlnLutThr 106
Db      583 TATATGTCGGCTTCAACGGAATACCTCACAGCAAAAGTAAACATTTATTTCAGAT 642
Qy      107 GlyAlaLys-----IleLeu 111
Db      643 GCTCTAAAGAGCAGTGAAGTGCACATGAGAAATTAATCTCGAACAACCATCTAGT 702
Qy      112 AspTyrIleProAspTyrAlaTyrIleValGlnLutThrLysAspValGlnSerLysVal 131
Db      703 GACTACCTTATGATTTTATG-----GTGATTCAGATTAAGAAAGAGAGAGGCGGCTG 759
Qy      132 ArgSerIleGlnHisValGlnLysSerValGlnLutProTyrLeuProLysTyrLysIleAspPro 151
Db      760 CTCACACTGAAGATCACCAACCAACATCAAGCGGTGACACCCACGCGAAAGCTTTCCT 819
Qy      152 GlnLeuPheThrLysGlyAlaSerThrLeuValLysAlaLeuAlaLeuAspThrLysGln 171
Db      820 TCCCTGAAGTTGCTGAATCCGACCCCATGTGGCC-----TGTATAGACACCCGCTG 873
Qy      172 AsnAsnLysGlnValGlnLeuArgGlyIleGlnGlnLysIleAlaGlnLutThrValAlaSerAsn 191
Db      874 AGCCAGAGTGGCAGTCAATCAGCTCCCTGAAAAGAGCCAGTCTCCCTGGGCTTGA 933
Qy      192 AspValHisTyrIleThrAlaLysProGlnLutThrLysValMetAsnAspValAlaArgGly 211
Db      934 TTCTGCAATGCAACAGAAAGCATTCAGTCAACATTCCTAGAGCCATTCTCTGCCAG 993

```



```

QY 212 lle-----VallysaiaapValAlaGlnSerSerTrgLyLeuTrgLyGlnGly 228
   :::::|||||
Db 994 GTTGGCCAGACATTCGACGGACATGTCCTTGG---CAGATGGAGATACACAGGTGCTMAAT 1050
QY 229 GlnIleValAlaValAlaAapTrhTrgLyLeuAapTrhTrgLyAArgAapSerSerMetHis 248
   |||||
Db 1051 GTCAGGGTTCGCTTTTGTATCTGGGCTC-----AGTGAGAGCAT 1092
QY 249 GluAlaPheAArgLyLySerIleThrAlaLeuTrgAlaLeuGlyAArgThraAapAlaAan 268
   |||||
Db 1093 CCACATTTCAAGAAAT-----GTGAGGAAAGAAACCAACTGACCAAT 1134
QY 269 -----AapThraAenGlyHieGlyThrHiValAlaGlySerValIleuGly 283
   |||||
Db 1135 GAGCGGACCCCTGGACATGGGCTGGGCCATGGACATGCTTGCTCAAGT---GTGATGGC 1191
QY 284 AenGlyAlaThraAenLyGlyMetAlaProGlnAlaAenLeu---ValPheGlnSerIle 302
   :::::|||||
Db 1192 AGCATGAGAGAGTGCACAGGATTTGGCCCATGATGCAGACCTGCACATCTTCAGGGCTT 1251
QY 303 MetAapSerSerLyLyLeuGlyLeuProSerAenLeuGlnThraAenPheSerGln 322
   :::::|||||
Db 1252 ACCAACATACAG-----GTGCTTACACGCTTGGTTTGGAT 1290
QY 323 AlaPheSerAlaGlyAlaAArgIleThraAenSerTrgLyAlaAlaValAenGlyAla 342
   |||||
Db 1291 GCCTTCAACTATGCCATCCCTAAAGAGATKAGCTTCTGACCTTAACCTTACCGGGGCT 1350
QY 343 TrpTrhThraAapSerAArgAenValAapAapTrpYr-----ValAArgLyAapAapMetTrh 360
   |||||
Db 1351 GACTTCATGATCAACCCCTTTTGTGACAGATGGAAATTAAACGCGAACATGTAATC 1410
QY 361 IleLeuPheAlaAlaGlyAenGlyAArgProAenGlyLyThrIleSerAlaProGlyTrh 380
   :::::|||||
Db 1411 ATGCTTTCGCTATTTGGCAATGAGACCTCTCTATGGACCTGTAATTAACCTCTGAT 1470
QY 381 AlaLyAapAlaIleThrValGlyAlaThrGluAenLeuAArgProSerPheGlySerTrp 400
   :::::|||||
Db 1471 CAGATGATGATGATGGAGTGGGTGGCATTGAC-----TTT 1506
QY 401 AlaAapAenIleAenHiValAlaGlnPheSerSerAArgLy----- 414
   |||||
Db 1507 GAAGACAAACATC-----GCCGCTTCTCTTCCAGGGGAAATGATCACTGGGAACATA 1557
QY 415 ProThraLyAapGlyAArgIleLyProAapValMetAlaProGlyThrTrpTrIleuSer 434
   |||||
Db 1558 CCGGAGAGCTATGCTGCTGTAAGCTTACATTTGTC-----ACCTATGGTCTGGA 1608
QY 435 AlaAArgSerSerLeuAlaProAapSerSerPheTrpAlaAenHiAapSerLySerAla 454
   |||||
Db 1609 GTGCGGGTTCGTGTGTGAAGGGGGCTGC-----CGT 1641
QY 455 TyrMetGlyLyThrSerMetAlaThrProIleValAlaGlyAapValAlaGlnLeuArg 474
   :::::|||||
Db 1642 GCACCTCAGGAGACAGTGTCCGCTCCAGTGTGCTGGGGCTGTCACTTGTATGTA 1701
QY 475 GluHisPheValIleAapAArgLyLyIleThrProLyProSerLeuLeuAlaIleu 494
   |||||
Db 1702 AGCACAGTACAGAACCGGAGCTAGTG-----AATCTGCCAGTGTGAAGAACCTTGG 1755
QY 495 IleAlaGlyAlaAlaAapValGlyLeuGlyTrpProAenGlyAap-----GlnGly 511
   |||||
Db 1756 ATAGCATCAGCCCGGAGACTT-----CCTGGTGTCAACATGTTTGAGCAAGGC 1803
QY 512 TrpGlyAArgValThraAapLyAapSerIleuAenVal-----AlaTrpValAenGlySer 529
   |||||
Db 1804 CATGCAAGTGTGATCTACTGAGACCTATCAAGATCCACAGAGCTATTAACCGGAGCG 1863
QY 530 SerAlaLeuSerThraSerGlnLyAalThrTrpThraAlaThraAlaGlyLyAapPro 549
   |||||
Db 1864 AGCCG-----AGTCTT 1875
QY 550 LeuLySerIleSerLeuVal-----TrpSerAapAlaProAlaSer 562

```

```

Db 1876 AACTACATGACCTGATCGATGTCCTTACATGTGGCCCTTACTGTCTCCAGCCATCTAC 1935
QY 563 ThraAlaSerValThraLeuValAapAapLeuAapLeuValIleThraAlaProAenGly 582
   |||||
Db 1936 TATGAGGAATGCCAACATGTTTAA-----GTCAACATCTCTCAATGCC 1980
QY 583 ThraGlyTrpValGlyAapAapPheSerAlaProPheAapAapTrp----- 598
   |||||
Db 1981 ATGGAGTTACAGAGAAATGTGTGATAGCT-----GACTGGCAGACCTATTGTA 2031
QY 599 ---AapGlyAArgAapAapValGluAenValPheIleAenSer-----ProGln 613
   :::::|||||
Db 2032 CCACGAATGAGACACATTTGAATGAGGCTTCTCTCACTCTCACTGTGTGGCTTGG 2091
QY 614 SerGlyThrTrpTrhIleGluVal 621
   |||||
Db 2092 TCAGGTTACCTTGCATCTCCATT 2115

```

```

RESULT 7
US-11-122-329-113
; Sequence 113, Application US/11122329
; Publication No. US20060019272A1
; GENERAL INFORMATION:
; APPLICANT: Geraci, Mark
; APPLICANT: Bull, Todd
; APPLICANT: Voelkel, Norbert
; TITLE OF INVENTION: Diagnosis of Disease and Monitoring of Therapy Using Gene
; FILE REFERENCE: 2848-54
; CURRENT APPLICATION NUMBER: US/11/122,329
; PRIOR FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/568,129
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 113
; LENGTH: 4338
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-122-329-113

```

```

Alignment Scores:
Pred. No.: 3,08e-14 Length: 4338
Score: 243.00 Matches: 140
Percent Simlarity: 38.6% Conservative: 89
Best Local Simlarity: 23.6% Mismatches: 224
Query Match: 7.4% Indels: 140
DB: 14 Gaps: 28

```

```

US-10-784-870-4 (1-639) x US-11-122-329-113 (1-4338)
QY 87 HisIleValGlnPheAenGlyProIleLeuGlnGluThraLyGlnLyVleuGluThrThr 106
   ::|||
Db 662 TATATGTGGCTTCAATGATACCTTACAGCCAAACCTAGAAATTATTTATTTCAAGT 721
QY 107 GlYAlaLy----- 111
   |||||
Db 722 GCCCTGAAGAGCAGTGAATGACAAATTGAGAAATTATCTCGAAACAATCCATCCAGT 781
QY 112 AapTrpIleProAapTrpAlaTrpIleValGlnTrpGluLyAapValGlnSerLyVal 131
   |||||
Db 782 GACTTACCTTGTGATTTTGAG---GTGATTCAGATTAAGAAAGAAAGCGGGGCTG 838
QY 132 ArgSerIleGluHisValGlnSerValGluProTrpLyLeuProLyTrpLyLySileAapPro 151
   :::::|||||
Db 839 CTAAACATTTGAATATCAAAACATCAAAAGGCTACGCCCCACAGAAAGTCTTTCGT 898
QY 152 GlnLeuPheThraLyGlyAlaSerThraLeuValIleValAlaLeuAapThraLyGln 171
   |||||
Db 899 TCCCTC-----AAGTATGCTGATGTGACCCCAAGTACCTTGCATATGAACCCGGTGG 952

```


OY 1728 AaaAaSuTlVgGluValGlInLeuRgLyLleGIguLuIleAlGlnTy-ValAlSerLen 191
 : : : :
 Db 953 ACCCGAAAGTGCCATCATCAAGTCCTCCCTGCCAAAGACCGACTCTCCCTGGGCTTGGC 1012
 OY 192 AapValHi.eTy.rIl.eThrAlalysPProGuTyTyVsValMeCaaAbApValAlaArgLy 211
 : : : :
 Db 1013 TTCTGGCATGTACGGGAAGGCAATTCAGACAACGGCTGTGTAAGGCATCCCGCGCAG 1072
 OY 212 Il.e-----VallyeAlaApValAlagInSeSerTy.rGlyLeuTy.rGlyGlnLy 228
 : : : : : : : :
 Db 1073 GTTGGCCAGACACTGACGCAATGCTCTGG---CAGATGGATTATACAGTCTAAT 1128
 OY 229 GlInLeValAlaValAlaApThrGlyLeuApThrGlyArgAsnAbpSerMetHis 248
 : : : : : : : :
 Db 1130 GTAAGAGTTGCTGTTTTTGA.CTGGGCTG-----AGCGAAGAAGAT 1177
 OY 249 GluaI.aPeArGeLyLy.eIl.eThrAlaleuTy.rAlaleuGlyArgThr.AsnAbnlaAn 268
 : : : : : : : :
 Db 1172 CCCCACTTCAAABAA------GTGAAGAGAGAACCAACTGACCAAC 1213
 OY 269 -----AepThrAnGly.yH.eGlyTy.rH.eValAlaGlySerVal.leuGly 283
 : : : : : : : :
 Db 1214 GAGCGAAGCTGAGCAATGGGTGGCCATGGCAATTCGTGGCAGGT--GTGATACC 1270
 OY 284 AnGlyAlaThr.AsnLyGlyMetAlaProGlnAlaenLeu--Val.PheGlnSerIle 302
 : : : : : : : :
 Db 1271 AGCATAGAGGAAGTGC.CAAAGATTTC.TCAAGTAGCAATTCACATTTTCAGGCTCTTT 1330
 OY 303 MetAapSerSerGlyLy.leuGlyLyLeuP.roSerAanLeuGlnTh.LeuPheSerLn 322
 : : : : : : : :
 Db 1331 ACCATATATCAG-----GATCTTACACATCTTGCTTTTGAC 1368
 OY 323 AlaPheSerAlaGlyAlaArgIl.eH.eThr.AnsErTy.rGlyAlaAlaVal.AsnGlyAla 342
 : : : : : : : :
 Db 1370 GCCTTCAACTATGCACTTTTAAAGAAATGCA.GCTTAAACCTGACATCGGGCGCCG 1429
 OY 343 Ty.rThrThr.AbpSerArGAsnVal.AspAerTy.r-----ValAlGly.AsnAbpMetThr 360
 : : : : : : : :
 Db 1430 GACTTCATGAGATCATCCGTTTGTTCACAAAGTGGGAAATTACACCTTAAACATGATATC 1489
 OY 361 IleLeuPhea.laAlaGly.AsnGlyUkrPro.AbnGlyLyThr.IleSerAla.ProGlyThr 380
 : : : : : : : :
 Db 1490 ATGCTTTCTGCTATTGGCAATGACGACCTCTTATGGCACTCTGAATTAACCTGCTGAT 1549
 OY 391 AlAluY.aAnAlaIleThrVal.GlyAlaThrGlu.SbnLeuAurProSer.PheGlySerTy.r 400
 : : : : : : : :
 Db 1550 CAATATGATGTGATTTGGAGTAGGCGGCAATGAC-----TTT 1585
 OY 401 AlaAbpAbnIl.eAbnH.eValAlaGlnPheSerSerArGly----- 414
 : : : : : : : :
 Db 1586 GAAGATTAATC------GCCCGCTTTCTTCAAGGGAATGATCACTGAGGAGCTA 1636
 OY 415 ProThrLySAbpGlyArgIl.eUkrPro.AbpValMetAla.ProGlyThrTy.rIl.eUser 434
 : : : : : : : :
 Db 1637 CGAGAGGCTACGGTGCAGATGAACCTACATGTC-----ACCTAAGTGTCTGC 1687
 OY 435 AlaArGerSerLeuAlaPro.AbpSerSerPheTrpAlaAbnH.eAbpSerLyTy.rAla 454
 : : : : : : : :
 Db 1668 GTGCGGGTTCGCGGTGAAGGGGGGCTG-----CGG 1720
 OY 455 TyrMetGlyLyThrSer.MeAlaThrPro.IleValAlaGly.AsnValAlaGlnLeuArg 474
 : : : : : : : :
 Db 1721 GCCCTCTCAGGGACCAAGTGGCTTCTCAAGGtGTGACAGGTGCTGACCTGTATAGT 1780
 OY 475 GluHIsPheValySAbnArgLyLleThrProLySProSerLeuLeuLyAlaLeu 494
 : : : : : : : :
 Db 1791 AGCACAGTCCAGAAACGTGAAGCTGTG-----AATCCCGCAGTATGACACAGCCCTG 1834
 OY 495 Il.eAlaGlyAlaAla.AspValAlaGlyLeuGlyTy.rPro.AbnGly.Asn-----GlnGly 511
 : : : : : : : :
 Db 1835 ATCGGCTCAGCCCGAGGCTC-----CCGGGGCTCACATGTTTGACAGACGC 1882
 OY 512 TrpGlyArGValThrLeu-----Asplys 519

DG

```
Bd      |::|::|::|::||  
        ||:  
Bd      1893 CAGCGCAACCTCATTGGTCGTACAACCCTATCATCTCTCAAACAGCTAACAGCCAAAGGCCA   1942  
  
Qy      SerLeuSenValAlaIYrVaLaSngLSeSerzAla-----LeuSerThirSer    535  
          ||:  
Dd      1943 AGTTTGAACCCCACACTACATGCATGCATGCATGCATGCATGCATGCATGCATGCATGCATGCCTC     2002  
          ||:  
Qy      GlnlybaalatrTYrrPheThrAlaThrAlaGlLYyrProleulyvlllSerLeuval    555  
          ||:  
Dd      2003 CAGCCCCATCTACTAT-----GAGAATAATGCCG-----                2029  
  
Qy      556 TrpSerapAlaPrOlaseThrThrAlaSerValnHreUvaJlaenApLeuaPlenu    575  
Dd      2030 -----ACAGTGTTTAATCTCACATCTCCAACGSCAACGSGCAAGGAGATC     2068  
          ||:  
Qy      576 -----VallletnralAProASngLythrArgTyrvAlGLyenadepHeSerAla    592  
          ||:  
Dd      2069 ACAGAGAAATYTtGaNGAtMAgCcCTGActGcGAcGccCTAttTtg-----       2110  
          ||:  
Qy      593 ProPheaPeaaenaanTPaSpGIYArGaAnaAnVaIgLuBanVaIPheileasSer---    611  
          ||:  
Dd      2111 -----CCACAGAACGAgAC-----AACTGAAATGTCgcCTTctTactCTGCTG     2155  
          ||:  
Qy      612 -----ProGInserGIYThrtYTrtllegluVal    621  
          ||:  
Dd      2156 GTCTTAGGCCTTgtGtcGgCTacTGCagCatTCCTCATT    2194  
  
RESULT 8  
US-11-047-380-4  
; Sequence 4, Application US/11047380  
; Publication No. US2006068406A1  
; GENERAL INFORMATION:  
; APPLICANT: AFPHOLTER, JOSEPH A.  
; APPLICANT: COX, ANTHONY  
; APPLICANT: NESS, JON E.  
; APPLICANT: CARR, BRIAN  
; TITLE OF INVENTION: SINGLE-STRAWDED NUCLEIC ACID TEMPLATE-MEDIATED  
FILE REFERENCE: 0165_551US  
CURRENT APPLICATION NUMBER: US/11/047,380  
PRIOR APPLICATION NUMBER: 09/721,507  
PRIOR FILING DATE: 2000-11-22  
PRIOR APPLICATION NUMBER: 09/656,549  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/185,244  
PRIOR FILING DATE: 2000-02-28  
PRIOR APPLICATION NUMBER: 60/185,915  
PRIOR FILING DATE: 2000-02-29  
PRIOR APPLICATION NUMBER: 60/186,247  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 60/186,482  
PRIOR FILING DATE: 2000-03-02  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 4  
LENGTH: 1180  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1)..(16)  
OTHER INFORMATION: a, c, t, g, other or unknown  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1163)..(1180)  
OTHER INFORMATION: a, c, t, g, other or unknown  
US-11-047-380-4
```

Alignment Scores:

```

Pred. No.: 7,18e-15 Length: 1180
Score: 242.50 Matches: 123
Percent Similarity: 37.3% Conservative: 55
Best Local Similarity: 25.8% Mismatches: 154
Query Match: 7.4% Indels: 145
DB: 11 Gaps: 19

US-10-784-870-4 (1-639) x US-11-047-380-4 (1-1180)

QY 1 MetArgLysLeuValPheLeuSerValLeuSerValAlaAlaIleLeuSerThrVal 20
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 17 GTGAGAGAGCAAAATTTGGATGACCTGTTGTCGTTAAGTTAATCTTACGANG 76

QY 21 AlaLeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIle 40
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 77 GCGTTACGACACATGTTCTGCGAGCGTCCGGAAGAAAGCAAGTCAAAAGAAATACATT 136

QY 41 GlnThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaIlePhe 60
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 137 -----GTGCGATT-----AAACAGCAATGAGTGCATGAGT 169

QY 61 LeuValGlnSerGluAsnValLysLeuLysGly-----LeuLeuLysLysGlu 78
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 170 TCCGCGCAAGAAAAGGATGTTATTTCTGAAAAGCGCGAAGGTTCAAAAGCAATTAAAG 229

QY 79 ThrValProAlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlu 98
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 230 TATGTTACCG----- 241

QY 99 ThrLysGlnLysLeuGlnThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAla 118
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 242 GCCGCGACAACTTGGATGAAAAGCTGTAAGAAATGAAAAGATCCGAGCGTTGCA 301

QY 119 TyrIleValGlnTyrGlnLysAspValGlnSerLysValArgSerIleGluHisValGlu 138
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 302 TATGTCGAAGAAAGATCAT-----ATTGCACATGATATATGCGCAA 340

QY 139 SerValGlnProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAla 158
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 341 TCTGTT-----CCTTAT----- 352

QY 159 SerThrLeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeu 178
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 352 ----- 352

QY 179 ArgGlyIleGlnGlnIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAla 198
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 353 ---GGCATTCTCAATT----- 367

QY 199 LysProGlnTyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspValAla 218
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 366 -----AAGCGCGCG---GCT 379

QY 219 GlnSerSerTyrGlyLeuTyrGlyGlnGlnIleValAlaValAlaAspThrGlyLeu 238
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 380 CTTCACTCTCAAGGCTACACAGGCTCTAAAGTAAAGTGGTATGACACGCGGAATT 439

QY 239 AspThrGlyArgAsnAspSerSerMetHisGlnAlaPheAspGlyLysIleThrAlaLeu 258
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 440 GACTCTTTCATCTGACTTAACGTC-----AGAGGC-----GAGACCAAGC 481

QY 259 TyrAlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAla 278
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 482 TTCGTAACCTTGAAACAAACCAACACAGACGCGCATCTTCACAGGTAACGATTAACC 541

QY 279 GlySerVal-----LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsn 296
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 542 GGTACGATTGCGCGCTCTTAATTAATCAATCGGTGTTGCGGCGTTAGCCCAAGCGCATCA 601

QY 297 LeuValPheGlnSerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeu 316
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 602 TTATATGCGAGTAAAGTCTGATTCAACAGCA-----ACGCGCCCAATATATGCTGAGTT 655

```

```

QY 317 GlnThrLysPheSerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyrGly 336
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 656 ATTTAAGCGCATTTAGTGGCCCATTTTCCAAATATGATGATGATGACATGAGCGCTTGGC 715

QY 337 AlaAlaValAsnGlyAlaTyrThrAspSerAspAsnValAspAspTyrValArgLys 356
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 716 GGACCTACTGCTG-----TCTACAGCGCTGAAACAGATCTGTTACAAAGCCGTTTCC 766

QY 357 AsnAspMetThrIleLeuPheAlaAlaGlyAsnGluArgProAsnGly-----GlyThr 374
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 767 AGCGGTATCGTGTGCTGCGCACGCCGAAACGAAGGTTATCCGGAAGCAAGACACA 826

QY 375 IleSerAlaProGlyThrAlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArg 394
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 827 GTGCGTACCTCGCAAAATATCTTCTACTTGTGACGATGGCG----- 871

QY 395 ProSerPheGlySerTyrAlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGly 414
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 872 -----GTAAACGACAGCAACCAAGACTTCTTCCAGCGCGAGT 913

QY 415 ProThrLysAspGlyArgIleLysProAspValMetAlaProGlyThrTyrIleLeuSer 434
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 914 TCT-----GAGCTTATGTGATGCTCTGCG-----GTGTC 946

QY 435 AlaArgSerSerLeuAlaProAspSerSerPheThrAlaAsnHisAspSerLysTyrAla 454
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 947 ATCCAAAGCACACTTCT-----GGAGGCACTTAACGC 979

QY 455 TyrMetGlyGlyThrSerMetAlaThrProIleValAlaGlyAsnValAla 471
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 980 GCTTATACGGAAGCTCATGCGACTCTCACTGCGGAGCAGCAGCGC 1030

RESULT 9
US-11-128-061-679
; Sequence 679, Application US/11128061
; Publication No. US2006003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William W.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; PRIOR FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 679
; LENGTH: 4199
; TYPE: DNA
; ORGANISM: Cricetus griseus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (130)..(150)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4177)..(4199)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-679

Alignment Scores:
Pred. No.: 3.76e-14 Length: 4199
Score: 242.00 Matches: 138
Percent Similarity: 38.1% Conservative: 86
Best Local Similarity: 23.5% Mismatches: 234

```

Query Match: 7.4% Indels: 130
DB: 14 Gaps: 27
US-10-784-870-4 (1-639) x US-11-128-061-679 (1-4199)

```

QY      87 HisIleValGlnPheAenGlyProIleLeuGlnGluThrIlySerGlnIlyLeuGlnThr 106
      552 TATATTGTGGCTTCAACGATATTCTTCAAGCCAAAGCTAGAACTATTATTCAAGT 611
QY      107 GYAlaIlyS-----11Leu 111
      612 GCTCTGAAAGGAGTGAAGTGAACACTGAGATTATCTCGGAACAACCCATCCAGT 671
QY      112 AspTyrIleProAspTyrAlaTyrIleValGluTyrGlnIlyAspValGlnIlySerVal 131
      672 GACTTACCTTGATGATTTTGAG--GTGATTCAGATTAAAGAAAGCAAGAGCCGGCTG 728
QY      132 ArgSerIleGlnIlyValGlnSerValGluProTyrIleuProIlySerIlyIleAspPro 151
      729 CTCACACTTGAGATCATCAAAACATCAAGCGGGTGCACCTCAAGCAAAAGCTTTCCGT 788
QY      152 GlnLeuPheThrIlySerGlnIlyAspThrLeuValIlyLeuAlaLeuAspThrIlySerGln 171
      789 TCCTTGAAGTTTCTGATCTGACCCCATTTGTCCA-----TGTATGAACTCCGGTGG 842
QY      172 AsnAspIlySerGlnIlyLeuArgGlyIleGlnGlnIleAlaGlnIlyValAlaSerAsn 191
      843 AGCCAAAGTGGCAGATCATCAACCCCTGAGAAAGCCGCTCTCCCTGGGCTCTGGA 902
QY      192 AspValHisTyrIleThrAlaIlyProGluTyrIlyValMetAsnAspValAlaArgGly 211
      903 TTCCTGATCATCAACAGAAAGATCAAGCCGGCATGTGCTGAGAGCCATTCCTCGACAG 962
QY      212 Ile-----ValIlyAlaAspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnIly 228
      963 GTTGCACCAACATTCAGAGGAGATGTGCTGTG---CAGATGGCAATACAGAGTGTCAAT 1019
QY      229 GlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHis 248
      1020 GTCCAGGTTGTGCTGTTTGTACTGGGCTC-----AGTGAAAGCAAT 1061
QY      249 GlnAlaPheArgGlyIlySerIleThrAlaLeuGlyIlyArgHisAsnAspIleAsn 268
      1062 CCACACTTCAGAAAT-----GTGAGAGAGAAACCACTGAGCAACAT 1103
QY      269 -----AspThrAsnGlyIlySerIleValIlyLeuValAlaGlySerValIleuGly 283
      1104 GAGCGGACCTGTGATGATGGCTGGCCCATGCGACATTTGTCCGACAGT---GTGATTGCC 1160
QY      284 AsnGlyAlaThrAsnIlySerMetAlaProGlnAlaAsnLeu---ValPheGlnSerIle 302
      1161 AGCATGAGGAGATGCCAGGATTTGGCCCATGATGAGCTGCAATCTTCCGGGCTCTTT 1220
QY      303 MetAspSerSerGlyIlyLeuGlyIlyLeuProSerAsnLeuGlnThrIleuPheSerGln 322
      1221 ACCAACAATCAG-----GTGCTTACCAATCTTGCTTTTGGAC 1259
QY      323 AlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyrGlyAlaAlaValAsnGlyAla 342
      1260 GCTTCAACTGATGCTCTTAAAGAAAGATTGATTTCTAAACCTTACATCGCGGGCTT 1319
QY      343 TyrThrThrAspSerArgAsnValAspAspTyr-----ValArgIlyAsnAspMetThr 360
      1320 GACTTCATGGATATCCCTTTGTTGCAAGGTGTGGAAATTAACACTTAACATGTAAATC 1379
QY      361 IleLeuPheAlaAlaGlyAsnGluAspProAsnGlyIlyThrIleSerAlaProGlyThr 380
      1380 ATGGTTTCTGCTATCGGCAATATGACCTTTTAAAGCACTCTGAATTAACCAAGCTGAT 1439
QY      381 AlaIlyAsnAlaIleThrValGlyAlaThrGlnAsnLeuArgProSerPheGlySerTyr 400
      1440 CAGATGATGTGATGTGAGTGGGTGCTGATTCAGC-----TTT 1475

```

```

QY      401 AlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGly-----414
      1476 GAGATTAATC-----GCCCGCTTTCTTCCAGGGGAATGACTTACCTGGAACTA 1526
QY      415 ProThrIlyAspGlyIlyArgIleIlyProAspValMetAlaProGlyThrTyrIleuSer 434
      1527 CCAGAGGCTATGATGTGGCGTGAACCTGACATTGTG-----ACCTATGGCGCGGA 1577
QY      435 AlaArgSerSerLeuAlaProAspSerSerPheThrAlaAsnHisAspSerIlyTyrAla 454
      1578 GTCCGGGCTTCCCGTGTGAAGGGGCTGC-----CGG 1610
QY      455 TyrMetGlyIlyThrSerMetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArg 474
      1611 GCACCTTCAGGAGACAGATGTGCTTCCCATGATGTTGCGGCTGTGACCTTGATGTA 1670
QY      475 GlnHisPheValIlyAsnArgGlyIleThrProIlyProSerLeuLeuValAlaLeu 494
      1671 AGCAGATGAGAAAGGGAGCTAGTG-----ATCTGCGCAGTGTGAAGCAAGCCCTG 1724
QY      495 IleAlaGlyAlaAlaAspValGlyLeuGlyTyrProAsnGlyAsn-----GlnGly 511
      1725 ATTCATCAAGCCCGGAGGCT-----CTGTGTAAATGTTGAGAGAGGC 1772
QY      512 TrpGlyArgValThrIleuAspIlySerLeuAsnVal-----AlaTyrValAsnGlnUser 529
      1773 CATGGCAGTGGATGTGCTGCGAGCTTACATGATCTCAGACGCTACAAACAGGCG 1832
QY      530 SerAlaLeuSerThrSerGlnIlyAlaThrTyrThrPheThrAlaThrAlaGlyIlyPro 549
      1833 AGCTTG-----AGCTCT 1844
QY      550 LeuIlyIleSerLeuVal-----TrpSerAspAlaProAlaSer 562
      1845 AGCTACATGACCTGACGTAGTGTCCCTACATGTGCTTACTGTTCTCAGCCATCTAC 1904
QY      563 ThrThrAlaSerValThrIleuValAsnAspLeuAspLeuValIleThrAlaProAsnGly 582
      1905 TATGAGAAATGCCCAACAAATGTTAAAT-----GTCCACATCTCAATGAGC 1949
QY      583 ThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAspTyr-----598
      1950 ATGGAGTGCAGAGAAATGTGATTAAGCT-----GAGTGGCGGCTTATTA 2000
QY      599 ---AspGlyArgAsnAsnValGluAsnValPheIleAsnSer-----ProGln 613
      2001 CCACAGAAATGAGACAAATTTGAAGTGGCTTCTCTACTCTCAGTGTATAGGCTTGG 2060
QY      614 SerGlyThrTyrThrIleGluVal 621
      2061 TCAGGCTACCTGGCCATCTCCAT 2084

```

RESULT 10
 US-11-128-049-679
 ; Sequence 679, Application US/11128049
 ; Publication No. US20060010513A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Melville, Mark W.
 ; APPLICANT: Charlebois, Timothy S.
 ; APPLICANT: Mounts, William M.
 ; APPLICANT: Ham, Louane E.
 ; APPLICANT: Sinacore, Martin S.
 ; APPLICANT: Leonard, Mark W.
 ; APPLICANT: Brown, Eugene L.
 ; APPLICANT: Miller, Christopher P.
 ; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
 ; FILE REFERENCE: 01997.027700
 ; CURRENT APPLICATION NUMBER: US/11/128,049
 ; CURRENT FILING DATE: 2005-05-11
 ; PRIOR APPLICATION NUMBER: US 60/570,425
 ; PRIOR FILING DATE: 2004-05-11
 ; NUMBER OF SEQ ID NOS: 7285

```

QY      303 MetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGln 322
      ::::
Db      1221 ACCAACACATAG-----GAGCTTCACACATCTGGTTTTCGAC 1255
      ::::
QY      323 AlaPheSerLeuGlyAlaArgLeuHisThrAsnSerTrpGlyAlaAlaValAsnGlyVala 342
      :|||:
Db      1260 GCTTTCACATGCCCATCTCTAAAGAAGATGTGTTCTTAACCTTACATCGCGGGCCCT 1311
      :|||:
QY      343 TyrThrThrAspSerArgAsnValAspAspTyr-----ValArgLyAsnAspMetThr 360
      :|||:
Db      1320 GACCTCATGATCATCTCTTTGTGACAAAGGNGNGGAATTTAACGCTAACATATGATATC 1377
      :|||:
QY      361 IleLeuPheAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThr 380

```

250 ATGGTTCCTGCTATGGCCTCTGAAATACCCAGCTGAT 143

Cy	381	AlLysAsnAlaIleTherValGlValaThrCluAnLeukrProSerPheGlySerTyr	400
	:: ::		:: ::
Db	1440	CAGATGGAGTGTCATTGCAGCTGGCGCATTAAC-----TTT	147
Cy	401	AlAspAsnIleAsnHisValaIacInPheSerArgLy-----	414

22 1370 GGAAGTAAACAC-----GCCGCTTCTCCAGGGAAGACTACCTGGACATA 152

101

[illegible]

1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396

455 TyrMetG]vG]vThrSerMetA]amhrProT]eVa]A]aG]vAaVvA]A]aG]rLcVvAa 474

db 1611 GCACTCTCAGGAGCACCAGTGTGCGCTTCCCCAGTGTGACGAGGACGTTACGTTA 167

475 GluHisPheValIysAsnArgGlyIleThrProLysProSerLeuLeuValAlaLeu 494

Db 1671 AGCACAGTGCAGAGCGGAGCTAGT-----AATCCTGCCAGTGTGAAGCAAGCCCTG 1724

495 ILEALGLYALAASPVALGLYLEUGLYTYRPROANGLYASN-----GLNGLY 511

DB 1/25 ATTCATCAGCCCGAGGCTT-----CCTGGTGTAAACATGTCGAGCAAGGC 1772

```

014 11017 gva1 m1 ucab p l y b s e i ucwa b n v d l -----Aldiy f va la b ng l u s e r   529
      |||          |||          |||          |||          |||          |||          |||          |||
      :::::       :::::       :::::       :::::       :::::       :::::       :::::       :::::

```

COY

[illegible]

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

[illegible]

563 ThrThrAlaSerValThrLeuValAsnAsnLeuValIleThrAlaProAsnGly 582

Db 1905 TATGAGGATGCCAACATTGTTAT-----GTCAACATCTCAATGCC 1949

QY 583 ThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTyr----- 598

Db 1950 ATGGAGTCA CAGGAGATTGTGATAAGCT-----GAGTGGCGGCCCTATTTA 2000

599 ---AspGlyArgAsnAsnValGluAsnValPheIleAsnSer-----ProGln 613

2001 CCACAGAAI GGAGCAACAATG AAGTGGCCCTCTCTCTACCTCCCTCAGTGATGGCCCTTGG 2060

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

US-10-510-386-217
; Sequence 217, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groch
; APPLICANT: Clausen, Steen Troels
; APPLICANT: Olsen, Peter Blarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 217
; LENGTH: 5296
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(4796)
US-10-510-386-217
Alignment Scores:
Pred. No.: 1,14e-13 Length: 5296
Score: 238.50 Matches: 163
Percent Similarity: 34.7% Conservative: 95
Best Local Similarity: 21.9% Mismatches: 248
Query Match: 7.3% Indels: 237
DB: 8 Gaps: 36
US-10-784-870-4 (1-639) x US-10-510-386-217 (1-5296)
QY 3 LysLysLysValPheLeuSerValLeuSerAlaAlaIle-----LeuSerThrVal 20
DB 516 AGAAAAAGCGCTTCCACCATTTTAAGCGTTTGAATCGGCTCGCTTCATGCC 575
QY 21 AlLeuLeuLeuLeuProSerAlaGlyAspAlaArgThrPheAspLeuAsp----- 36
DB 576 GCGGTTTCAGATGCCGCGGGAAGCCCGCACTTACCTTCAGAAAGAAAGAGCCGCA 635
QY 37 PheLysGlyIleGlnThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGly 56
DB 636 GGGAAAGGGAATTTTCAAAACCTCTTGTCAACAGTTCAAAAAGAAAGATCAA----- 689
QY 57 AlaAlaAlaPheLeuValGluSerGluAsnValLysLeuLeuLysGlyLeuLysVal 76
DB 690 ---GTCACTTTTGTGATTAAGTGAAGATCAAGTCATACACAAAGAGTGGCCAAACAG 746
QY 77 LeuGluThrValProAlaAsnAsnLysLeuHisIleVal-----GlnPheAsn 92
DB 747 GCGCAAAAGAACCGAAGAAAGTCCGTGAGCGCTCTAAACAGAAATCAAAAACGC 806
QY 93 GlyProIleLeuGluGluThrLysGlnLysLeuGluThrThrGlyAlaLysIleLeuAsp 112
DB 807 TCGGCTGTGTTCCTTAAAGATCAAGCGGATGATCGCAAAAGCATCTGAAACGT 866
QY 113 TyrIleProAspTyrAlaTyrIleValGluTyrGluGlyAspValGlnSerLysValArg 132
DB 867 TATTTAAAAAAGCAG-----GAAAAACAGGGGACGTAAAA---AAATTCAGA 911
QY 133 Ser-----IleGluHisVal 137
DB 912 TCTATATATATGTCAACGCGCATGCGCTCATGCGAAGAGATCATGGAACAAGTA 971
QY 138 GluSer-----ValGluProTyrLysProLysTyrLysIleAspProGlnLeuPhe 154
DB 972 GCGGCTTTTCCGAAGAGAAAAAGTCTTCCGAACGAAAA----- 1013
QY 155 ThrLysGlyValAspThrLeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLys 174
DB 1014 ---AGGCAGCTGATCAACCGACTAAACGCTGCTGTGAAA 1049

QY 175 GluValGlnLeuArgGlyIleGluGlnIleAlaGlnTyrValAlaSerAsnAspValHis 194
DB 1050 AATCAACACGACCCAGATGAAGAAAGATCAATG-----AATATCAAT 1094
QY 195 TyrIleThrAlaLysProGluTyrLysValMetAsnAspValAlaArgGlyIleValLys 214
DB 1095 CAGTCAACGACCAAAAGCTGGAATTC----- 124
QY 215 AlaAspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAla 234
DB 1125 -----GATATGACGATCGGACGCGGATGTCGCTCATTT 1160
QY 235 AspThrGlyLysLeuAspThrGlyLysAsnAspSerSerMetHisGluAlaPheArgGly--- 253
DB 1161 GATACGGCGGTGAGTGG-----GATCATCCGCTTGAAGAAATATCCGCGGATAT 1214
QY 254 -----LysIleThrAlaLeuTyrAlaLeuGlyArg 263
DB 1215 GACCCGACGATCCCATCAAGCTTACCGAATTCAGTTGTATGATGCGGTTTCAGGC 1274
QY 264 ThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGly 283
DB 1275 GCTTGTGAACCATTTGACATCTGAGCAGCAGCAGCATGTAAACCGGACGATGTCGCGC 1334
QY 284 -----AsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGln 300
DB 1335 TCAGAGCTTCACGCGCCAAACCAATTCGGCGTGGCGCGGATGATGATTCGTG 1394
QY 301 SerIleMetAspSerSerGlyLysGlyGlyLeuProSerAsnLeuGlnThrLeuPhe 320
DB 1395 AAAGCTTTTCAGATGACGCGGATACGAC----- 1424
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIle----- 330
DB 1425 GAAGACCTTATTCAGAGGTGATGATTTTGGCCGCAAAAGATAAAGACGCAACCG 1484
QY 331 His-----ThrAsnSerTrp-----GlyAlaAlaValAsn 340
DB 1485 CACCTGAAATGCGCGCGATGTCATTAATCTCATGATGTCAGAGGAGACCGGATGAT 1544
QY 341 GlyAlaIleTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThr 360
DB 1545 GAGTTTAC-----AGGATATGTAAGACGTGGAGCTGCGGCATCTTT 1592
QY 361 IleLeuPheAlaIleGlyAsn-----GluArgProAsnGly---GlyThrIleSer 376
DB 1593 CCGAGTTCCTCGCGGGAACGTGATTCGCAATCCGGAAGACCTGCTTCGATTCGC 1652
QY 377 AlaProGlyThrAlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSer 396
DB 1653 AATCCCGCAAACTACCTGAAGCGTTTCGCAACCGGACGCA----- 1694
QY 397 PheGlySerTyrAlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThr 416
DB 1695 -----ACATCATTAATCGCTTACCGGATTTTCACTTCAAGGCTCTCC 1739
QY 417 LysAspGlyArgIleLysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArg 436
DB 1740 CCGTATGATGACGATTAAGCTGAATATCGGCGCTGCTGTCACATT-----CGC 1790
QY 437 SerSerLeuAlaProAspSerSerPheTrpAlaAsnHisAspSerLysTyrAlaTyrMet 456
DB 1791 TCATCTGTG---CTGGAAGCGGCTTCAAGACGCTGGAC----- 1829
QY 457 GlyGlyThrSerMetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHis 476
DB 1830 ---GGCACTTCATGCGTGGCGCGATGTCGAGCGCGCGCTCTTATCAAGCAAGCC 1886
QY 477 Phe-----Val 478
DB 1887 GATTCGTGATCATCTTGTATGATGACGAAAGATTTTGTATGAAACTGACACCGCTC 1946

QY 479 LysAsnArgGlyLeuThrProLysProSer-----LeuLeuLysAla 492
DB 1947 ACGAGACGCAATTATTCAGATCGCGAACACGATACGCCACGAGCTCGGAACGTG 2006
QY 493 AlaLeuIleAlaGlyAlaAlaAspValGlyLeuGlyTyrProAsnGlyAsnGlyTyr 512
DB 2007 TTGACGACGAGTCCGCTGTGACAGACGGTTGGCGAGACGGAGACAAAGTC----- 2060
QY 513 GlyArgValThrLeuAspLysSer-----LeuAsnValAlaTyrValAsnGlySer 529
DB 2061 GGAAGAACGACGACAAAGATCGCGCCGCTTAACCATCAAGATGACCCGAA--- 2117
QY 530 SerAlaLeuSerThrSerGlnLysAlaThrTyrThrPheThrAlaThrAlaGlyLysPro 549
DB 2118 -----ATTACTCAGGCTCGAAGC-----CCT 2141
QY 550 LeuLys-----IleSerLeuVal----- 555
DB 2142 CTTAAAGCTGATGTAGAAGACGATGTCAGCATCATTCGGTCAAAATTGAGCTATAAAC 2201
QY 556 -----TrpSerAlaProAlaSerThrThrAlaSerValThrLeuValAsn 571
DB 2202 GATGACGCGCATTTGGAGACGATCGCTGCCAACAACTCCGCGACATATAAAAGGA 2261
QY 572 AspLeuAspLeuValIleThrAlaProAsnGlyThr-----ArgTyrValGly 587
DB 2262 ACCTACGAGCGGCTGACTGTTCTTCTGAAGGAGAAACCTTCTCTACAGTGGATGTC 2321
QY 588 AsnAspPheSerAlaProPheAsnAsnMetTrpAspGlyArgAsnAsnValGluAsnVal 607
DB 2322 ACCGATTTCT-----GAGGCAACAAACGGAATCC--- 2351
QY 608 PheIleAsnSerProGlnSerGlyThrTyrThrIleGluValGlnAlaTyrAsnValPro 627
DB 2352 -----AAAGTGTACGAAGTCCG 2369
QY 628 ValGlyPro 630
DB 2370 ATTTCCTCCG 2378
RESULT 12
US-11-156-062-13
; Sequence 13, Application US/1156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTION VARIANTS, WITH IMPROVED PHENYLDOLASE ACTIVITY
; FILE REFERENCE: HEK-0134 / H698
; CURRENT APPLICATION NUMBER: US/11/156,062
; PRIOR FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
; LENGTH: 824
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Variant T58A/L216W of Subtilisin Carlsberg
; NAME/KEY: CDS
; LOCATION: (1)..(822)
US-11-156-062-13
Alignment Scores: 2.09e-13 Length: 824
Pred. No.:

Score: 226.50
Percent Similarity: 43.1%
Best Local Similarity: 30.6%
Query Match: 6.9%
DB: 14
Gaps: 11
US-10-784-870-4 (1-639) x US-11-156-062-13 (1-824)
QY 212 IleValAlaAspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnIleVal 231
DB 28 CTCATTAAGCGGACAAAGTCCAGGCTCAA---GGCTTAAGGACGCAATGTAAAGTA 84
QY 232 AlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSerMetIleGluAlaPhe 251
DB 85 GCGTCTCGGATACAGAAATCCAGCTTCATCCGACTTGAACTTATCCGCGGA--- 141
QY 252 ArgGlyLysIleThrAlaLeuTyrAlaLeuGlyArgThrAsnAlaAsnMetThrAsn 271
DB 142 -----GCAAGCTTTGGCTGGCGAAGCTTATTAAGCC---AGCGCAAC 183
QY 272 GlyHISGlyThrHISValAlaGlySerVal-----LeuGlyAsnGlyAlaThrAsnLys 289
DB 184 GGAACGCGACACATGTTCCGCTACAGTACGTCGCTTGACACATCAACGGGTGTTTA 243
QY 290 GlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerGlyLeu 309
DB 244 GCGGTTCGCCAAGCGTATCTGTACGCGGTAAAGTACTGAATTCAGCCGA----- 297
QY 310 GlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAlaArg 329
DB 298 ACGGATCATACAGCGGCAATTGTAAACGGAATCGAGTGGCGCAACAACGCGATGAT 357
QY 330 IleHISThrAsnSerTrpGlyAlaAlaValAlaGlyAlaTyrThrAspSerArgAsn 349
DB 358 GTTATCAATATGAGCTTGGCGGACGA-----TAGGCTTCACACGCGATGAACAG 408
QY 350 ValAspAspTyrValArgLysAsnAspMetThrIleLeuPheAlaAlaGlyAsnGlyArg 369
DB 409 GCAATCGACATGATATGCAAGAGGGTCTGTTGATGCTGCAGCAGGGAACGCGGA 468
QY 370 ProAsnGly-----GlyThrIleSerAlaProGlyThrAlaLysAsnAlaIleThrVal 387
DB 469 TCTTCAGAAACAGCAATATCAATTTGGCTATCTGGGAATACGATTCGTCTCTGTT 528
QY 388 GlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsnIleVal 407
DB 529 GGTGCG-----GTAGACTCTTAACAGCAAGA 555
QY 408 AlaGlnPheSerSerArgGlyProThrLysAspGlyArgIleLysProAspValMetAla 427
DB 556 GCTTCATTTTCAGGCTCGA-----GCAGACCTTGAAGTCAATGGCT 597
QY 428 ProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerPheTrpAla 447
DB 598 CTTGCGCAGCGGCTATACAGCACTTAACCAACG----- 630
QY 448 AsnHISAspSerLysTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIleValAla 467
DB 631 -----AACACTTATGCAACATGGAACGGAAGTCAATGCTTCTCATATAG 681
QY 468 GlyAsnValAla-----GlnLeuArgGlnHISPheValLysAsn 480
DB 682 GAGGACGACGCTTTGATCTTGCAAAACATCGAACTTTACGTTCAACAAATCCGCAAC 741
QY 481 Arg 481
DB 742 CGT 744
RESULT 13
US-11-065-943-59
; Sequence 59, Application US/11065943
; Publication No. US20050250131A1
; GENERAL INFORMATION:

APPLICANT: JESTIN, JEAN-LUC
APPLICANT: VICHIER-GUERRE, SOPHIE
APPLICANT: PERRIS, STEPHANE
TITLE OF INVENTION: METHODS FOR OBTAINING THERMOSTABLE ENZYMES, DNA POLYMERASE I
TITLE OF INVENTION: VARIANTS FROM THERMUS AQUATICUS HAVING NEW CATALYTIC ACTIVITIES,
TITLE OF INVENTION: METHODS FOR OBTAINING THE SAME, AND APPLICATIONS OF THE SAME
FILE REFERENCE: 266426USOXIP
CURRENT APPLICATION NUMBER: US/11/065,943
PRIOR FILING DATE: 2005-02-25
PRIOR APPLICATION NUMBER: US 10/787,219
NUMBER OF SEQ ID NOS: 106
SOFTWARE: Patent in version 3.3
SEQ ID NO 59
LENGTH: 825
TYPE: DNA
ORGANISM: Bacillus amyloliquefaciens
US-11-065-943-59

Alignment Scores:

Pred. No.:	2,09e-13	Length:	825
Score:	226.50	Matches:	82
Percent Similarity:	39.4%	Conservative:	29
Best Local Similarity:	29.1%	Mismatches:	88
Query Match:	6.9%	Indels:	83
DB:	14	Gaps:	10

US-10-784-870-4 (1-639) x US-11-065-943-59 (1-825)

213 ValIyAlaAspValAlaGlnSerSerTyGlyLeuTyGlyGlnGlyGlnIleValAla 232
31 ATTAAAGCCCT---GCTTGCACCTCTCAAGGCTACAGTCAATGTAAGTAAAGCC 87
233 ValAlaAspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGluAlaPheArg 252
88 GTTATCGACAGCGGTATCAT---TCTTCATCTGATTAAG 129
253 GlyLysIleThrAlaLeuTyAlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGly 272
130 GTAGCAGCGCGAGCCGAGCATGGTCTCTGAAACAAATCTTCCAGACACAACTCT 189
273 HisGlyThrHisValAlaGlySerVal-----LeuGlyAsnGlyAlaThrAsnLysGly 290
190 CACGGAATCTCACGCTTCCGCGACAGTTGGCGCTTAAATACATCGGTGATTAGGC 249
291 MetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerGlyGlyLeuGly 310
250 GTTGGCGCAGCGCATCTTACGCTGTAAGATTCTCGGTCGACGCT----- 300
311 GlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAlaArgIle 330
300 ----- 300
331 HisThrAsnSerTrpGlyAlaAlaValaAsnGlyAlaTyThrThrAspSerArgAsnVal 350
301 ---TCCGGCATAATACGCTGATCATTAACGAAATGAGTGGCGCATCGAAACAAATATG 357
351 Asp----- 351
358 GACGTAATTAAACATGAGCCTCGCGGAGCCTTCTGCTGCTGCTTTAAAGCGGAGTT 417
352 AspTyValAlaGlyAsnAspMetThrIleLeuPheAlaAlaGlyAsnGluArgProAsn 371
418 GATTAAGCGGTGATCCGCGCTCGTACGTCGTCGCGCAGCCGCTAACGAAAGCACCTTCC 477
372 Gly-----GlyThrIleSerAlaProGlyThrAlaLysAsnAlaIleThrValGlyAla 389
478 GGAAGCTCAAGCAGCATGGGCTACCTGGAATACCTTCTGTCATGACAGAGCGCT 537
390 ThrGluAsnLeuArgProSerPheGlySerTyAlaAspAsnIleAsnHisValAlaGln 409
538 -----GTTGACAGCAGCAACCAAGAGCATCT 564

410 PheSerSerArgGlyProThrLysAspGlyArgIleTyProAspValMetAlaProGly 429
565 TTCGAGCGGTAGACCT-----GAGCTTATGTCATGACGACCTGGC 606
430 ThrTyIleLeuSerAlaArgSerSerLeuAlaProAspSerPheThrAlaAsnHis 449
607 -----GTACTATCCAAAGCACGCTTCC----- 630
450 AspSerLysTyAlaArgMetGlyIleThrSerMetAlaThrProIleValAlaGlyAsn 469
631 GGAACAAATATACGGCGCGTCAACAGGTACATGATGCTTCGACGTTGCCGAGCG 690
470 ValAla 471
691 GCTGCT 696

RESULT 14

US-11-156-062-11

Sequence 11, Application US/11156062

Publication No. US20050281773A1

GENERAL INFORMATION:

APPLICANT: Wieland, Susanne

APPLICANT: Polanyi-Bald, Laura

APPLICANT: Prueser, Inken

APPLICANT: Stehr, Regina

APPLICANT: Maurer, Karl-Heinz

TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY

FILE REFERENCE: HENK-0134 / H5698

CURRENT APPLICATION NUMBER: US/11/156,062

PRIOR FILING DATE: 2005-06-17

PRIOR APPLICATION NUMBER: PCT/EP2003/014127

PRIOR FILING DATE: 2003-12-20

PRIOR APPLICATION NUMBER: DE 102 60 903.9

PRIOR FILING DATE: 2002-12-20

NUMBER OF SEQ ID NOS: 24

SOFTWARE: Patent in version 3.3

SEQ ID NO 11

LENGTH: 824

TYPE: DNA

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: Variant T58A/L89S/N96D/L216W/N217D of Subtilisin Carlsberg

NAME/KEY: CDS

LOCATION: (1)..(822)

US-11-156-062-11

Alignment Scores:

Pred. No.:	2,65e-13	Length:	824
Score:	225.50 <td>Matches:</td> <td>86</td>	Matches:	86
Percent Similarity:	42.7% <td>Conservative:</td> <td>34</td>	Conservative:	34
Best Local Similarity:	30.6% <td>Mismatches:</td> <td>108</td>	Mismatches:	108
Query Match:	6.9% <td>Indels:</td> <td>53</td>	Indels:	53
DB:	14	Gaps:	11

US-10-784-870-4 (1-639) x US-11-156-062-11 (1-824)

212 IleValIyAlaAspValAlaGlnSerSerTyGlyLeuTyGlyGlnGlyGlnIleVal 231
28 CTCATTAAAGCGGACAAATGCAAGCTCA---GCTTTAAGGAGCGGAATGTAAGTA 84
232 AlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGluAlaPhe 251
85 GCCGCTCTGATACAGAAATCCAGGCTTCTCATCCGAACTGAACGTAGTCGCGGGA--- 141
252 ArgGlyLysIleThrAlaLeuTyAlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsn 271
142 -----GCAAGCTTTGTGCTGCGGAGCTTATTAAGCC---GACGCGAAC 183
272 GlyHisGlyThrHisValAlaGlySerVal-----LeuGlyAsnGlyAlaThrAsnLys 289
184 GGAACGCGCAGCATGTTGCCGTTACAGTACGTCGCGCTGACATATCAACGGGTGTTTA 243

Oy	290	GlyMetAlaProGlnAlaIleuValPheGlnSerIleMetLeuSerSerGlyGlyLeu	309
Db	244	GCGCTTGGCCCAAGCGTATCTCTGTCGACGGCTTAAAGTACTGATTCACGGCA-----	297
Oy	310	GlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAlaArg	329
Db	298	AGCGGATCATACAGCGCGCATTTGTAACGGCAATCGATGGCGGCACCAACAGCGCATATGAT	357
Oy	330	IleHisThrAsnSerTrpGlyAlaAlaValAsnGlyAlaArgThrThrAspSerArgAsn	349
Db	358	GTATATCAATATGACGCTTGGGGAGCA-----TCAGGCTTCGACAGCATGAACAG	408
Oy	350	ValAspAspTrpValArgIlyAsnAspMetThrIleLeuPheAlaAlaGlyAsnGlyArg	365
Db	409	GCAATCGAACAATGCATATATCGAGAGGGGTCGTCTGTGTACTGACAGGAGAACAGCGGA	468
Oy	370	ProAsnGly-----GlyThrIleSerAlaProGlyThrAlaIysAsnAlaIleThrVal	387
Db	469	TCTTCAGGAAACACGAATACAAATTTGGCTATCTCGGAAATACGATTTCTGCATGCGTGT	528
Oy	388	GlyAlaThrGluAsnLeuArgProSerPheGlySerTyrrAlaAsnIleAsnIleAsnIleVal	407
Db	529	GGTGG-----GTACACTTAAACACACACAGA	555
Oy	408	AlaGlnPheSerSerArgGlyProThrIlyAspGlyArgIleIysProAspValMetAla	427
Db	556	GCTCATTTTCCAGCGCTCGGA-----GCAAGAGCTTGAAGTCATAGGCT	597
Oy	428	ProGlyThrTyrrIleLeuSerAlaArgSerSerLeuAlaProAspSerSerPheTrpAla	447
Db	598	CCTGGCGCAGCGGCTATACAGCACTTACCCACAG-----	630
Oy	448	AsnHisAspSerLysTyrrAlaArgMetGlyGlyThrSerMetAlaThrProIleValAla	467
Db	631	-----AAACACTATGACCAATGAAACGAACGTCATATGCTTCTCTCATGTAGCG	681
Oy	468	GlyAsnValAla-----GlnLeuArgGluHisPheValIysAsn	480
Db	682	GGAGCAGCAGCCTTGTGATCTTGTCAAAACATCCGAACCTTTCAGCTTCAACATGCCGAC	741
Oy	481	Arg	481
Db	742	CGT	744
RESULT 15			
US-11-008-331-1/c			
; Sequence 1, Application US/11008331			
; Publication No. US20050244925A1			
; GENERAL INFORMATION:			
; APPLICATION: Genencor International, Inc.			
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms			
; FILE REFERENCE: GC9386-PCT			
; CURRENT APPLICATION NUMBER: US/11/008,331			
; CURRENT FILING DATE: 2004-12-08			
; PRIOR APPLICATION NUMBER: EP9719637.2			
; PRIOR FILING DATE: 1997-09-15			
; NUMBER OF SEQ ID NOS: 6			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 1			
; LENGTH: 2760			
; TYPE: DNA			
; ORGANISM: Bacillus subtilis			
US-11-008-331-1			
Alignment Scores:			
Pred. No.: 2,07e-12 Length: 2760			
Score: 223.00 Matches: 110			
Percent Similarity: 39.3% Conservative: 44			
Best Local Similarity: 28.1% Mismatches: 141			
Query Match: 6.8% Indels: 98			
Gaps: 13			

Qy 503 leuGlyTyrProAengIyAengIngIyTrpGlyArgValThrLeuAapLySerLeuAsn 522
Db 1684 -----TGGAAAAGGTTAATCAACGTACAAAGCAGCTGC 1652
Qy 523 ValAlaTyrValAengIuSerSerAlaLeuSerThr 534
Db 1651 ACAATATAGTAAAAAGAGCAGGTCTCCATACC 1616

Search completed: April 8, 2006, 08:55:20
Job time : 2743 secs

18:

18:

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 00:56:02 ; Search time 9420.85 Seconds
(without alignments)
11602.984 Million cell updates/sec

Title: US-10-784-870-5
Perfect score: 1923
Sequence: 1 atgagaagaagaagaagc.....cgttcgcaatgtgaattaa 1923

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_in:.*
3: gb_env:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pr:.*
9: gb_pro:.*
10: gb_srs:.*
11: gb_sy:.*
12: gb_un:.*
13: gb_vl:.*
14: gb_hlg:.*
15: gb_pl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1923	100.0	1923	1	AB051423	AB051423 Bacillus
2	1923	100.0	1923	6	CQ888285	CQ888285 Sequence
3	1923	100.0	1923	6	CQ889309	CQ889309 Sequence
4	1923	100.0	1923	6	AR368117	AR368117 Sequence
5	1923	100.0	1923	6	AR562465	AR562465 Sequence
6	1919.8	99.8	1923	6	AR368118	AR368118 Sequence
7	1919.8	99.8	1923	6	AR562466	AR562466 Sequence
8	1887.8	98.2	1923	1	AB084155	AB084155 Bacillus
9	1477.2	76.8	1920	1	AB046403	AB046403 Bacillus
10	1477.2	76.8	1920	6	AR562464	AR562464 Sequence
11	1477.2	76.8	1920	6	AR562464	AR562464 Sequence
12	1305	67.9	1305	6	AX839476	AX839476 Sequence
13	1166.4	60.7	3003	6	AR069954	AR069954 Sequence
14	1166.4	60.7	3003	6	BD062155	BD062155 Nucleic a
15	936.2	48.7	2218	6	E03808	E03808 DNA encodin
16	854.6	44.4	1302	1	AB046406	AB046406 Bacillus
17	740.4	38.5	1299	1	AB046405	AB046405 Bacillus
18	738.8	38.4	1299	1	AB046402	AB046402 Bacillus

19	737.2	38.3	1299	1	AB046404	AB046404 Bacillus
20	81	4.2	346148	1	BX942652	BX942652 Bdeliowib
21	54	2.8	7218	6	166494	166494 Sequence 14
22	53.6	2.6	2000	6	AX655393	AX655393 Sequence
23	50	2.5	2000	6	AX655393	AX655393 Sequence
24	48	2.5	54	6	AX839494	AX839494 Sequence
25	47.2	2.5	5890	2	DDU60086	DDU60086 Sequence
26	46.6	2.4	1239	1	AF305633	AF305633 Thermoma
27	46.6	2.4	1239	1	AY028704	AY028704 Thermoma
28	45.2	2.4	40499	8	AC134821	AC134821 Homo sapi
29	45.2	2.4	168242	8	AC145035	AC145035 Homo sapi
30	45.2	2.4	168242	8	AC140830	AC140830 Homo sapi
31	45.2	2.4	173275	14	AC145031	AC145031 Homo sapi
32	45	2.3	415	6	AX566129	AX566129 Sequence
33	45	2.3	12295	6	AE013026	AE013026 Thermoma
34	45	2.3	146831	8	HS523C21	HS523C21 Human DNA
35	44.8	2.3	125020	8	AF429315	AF429315 Homo sapi
36	44.6	2.3	154712	14	AC139734	AC139734 Sus croce
37	44	2.3	176462	8	AC159827	AC159827 Pan trogl
38	43.6	2.3	132384	8	AC116347	AC116347 Homo sapi
39	43.6	2.3	168430	14	AC145037	AC145037 Homo sapi
40	43.6	2.3	175559	14	AC145101	AC145101 Homo sapi
41	43.4	2.3	110000	15	CR382122_08	CR382122_08 Contamination (9 of
42	43.2	2.2	211647	14	AC114167	AC114167 Rattus no
43	43	2.2	250541	1	AE017004	AE017004 Bacillus
44	42.8	2.2	136613	14	AC164927	AC164927 Ginglymos
45	42.4	2.2	834	6	BD110490	BD110490 EST and e

ALIGNMENTS

RESULT 1	AB051423	1923 bp	DNA	linear	BCT 10-MAY-2002
LOCUS	Bacillus sp. KP43	PROF gene for protease, complete cds.			
DEFINITION	AB051423				
ACCESSION	AB051423.2	GI:20521154			
VERSION	AB051423.2	GI:20521154			
KEYWORDS	Bacillus sp. KSM-KP43				
SOURCE	Bacillus sp. KSM-KP43				
ORGANISM	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
REFERENCE	1	Itoh, S. and Saeki, K.			
TITLE	new protease				
JOURNAL	Unpublished				
REFERENCE	2.	(bases 1 to 1923)			
AUTHORS	Saeki, K.				
TITLE	Direct Submision				
JOURNAL	Submitted (21-NOV-2000) Katsumasa Saeki, KAO CORPORATION;				
	2606,AKABANE,ICHIRAIMACHI, HAGA, TOCHIGI 321-3486, Japan				
	(E-mail:3871850@kaietnet.kao.co.jp, Tel:81285687471 (ex.7471),				
	Fax:81285687403)				
COMMENT	On May 9, 2002 this sequence version replaced gi:14164344.				
FEATURES	Location/Qualifiers				
Source	1..1923				
	/organism="Bacillus sp. KSM-KP43"				
	/mol_type="genomic DNA"				
	/strain="KP43"				
	/db_xref="taxon:109322"				
gene	1..1923				
	/gene="PROF"				
CDS	1..1923				
	/gene="PROF"				
	/codon_start=1				
	/transl_table=1				
	/product="protease"				
	/protein_id="BAB5674.2"				
	/db_xref="GI:20521155"				
	/translation="MRKKKVLSTLSAAALSTVALSNPSAGARPNFLDPKGIQTT				
	TDKPGSKGQGTGAALFVSESNVLPKGLKKLETPANNLGLIIQFNGPIETGKQ				
	OLEKQKALIDYIPYAYIVVEYGVKSTSTIHHVSEPLPIYRIDPOLPTGAS				
	ELVKAVLDTKQKKEVQLRGIEQLAQPAISNDVLTITPAKPKYKMDVAKGIVADV				

AOSSYGLVGGGQIVAVADTGLDTRNDSSWHEAFRGKITTALVALGRNNANDNIGHT
HVAAGSYLNGSTNKKMAPOANLYPOSIMSGGIGLBSNLOLFEQOASARAHITN
SWGAYNGATYDSDSRVDDYVRKNDITILPAAENBPNGGITLSABETAKOALTGATN
NUNPSFGSTADNINHVAQFSSRGPTDGRKPDVMAAGPTLISABETAKOALTGATN
DSKRYAMGGTSMATPIVAGNVADLRHFPVNRGITPSPILAKRALJAGADISFPMNH
GNOGWRVTLDLKSLNVAYNVESLSSTSKATYSFATAGKPKLKSILWMSDAPASTA
SYLVNDLDELITAPANGOTYVNDFTSPYNDMDGRNNEVNIINAPSGTYTIEVQA
YVNPVGQPFPSLAIVN"

ORIGIN

Query Match	100.0%; Score 1923; DB 1; Length 1923;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGACGATTTGTGACT 60	
DB	1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGACGATTTGTGACT 60	
QY	61 GTTGCTTAAGTATCATCTGAGGTGTCAGAGAAATTTGATCTGATTTCAAGGA 120	
DB	61 GTTGCTTAAGTATCATCTGAGGTGTCAGAGAAATTTGATCTGATTTCAAGGA 120	
QY	121 ATTGAGACAACTGATGCTAAAGTTTCTCAAGCAGGGGAGACGTGCTGCT 180	
DB	121 ATTGAGACAACTGATGCTAAAGTTTCTCAAGCAGGGGAGACGTGCTGCT 180	
QY	181 TTTCTGCTGAATCTGAAATGTGAACCTCCAAAGGTTTGAGAGAAAGCTTGAACA 240	
DB	181 TTTCTGCTGAATCTGAAATGTGAACCTCCAAAGGTTTGAGAGAAAGCTTGAACA 240	
QY	241 GTCCCGGCAATATTAACCTCATTTATTCATTTAGATGACCAATTTAGAGAAACA 300	
DB	241 GTCCCGGCAATATTAACCTCATTTATTCATTTAGATGACCAATTTAGAGAAACA 300	
QY	301 AAACAGCAGCTGAAAGAAAGGGGCAAGATTTCTGACTACATCTGATTAAGTTAC 360	
DB	301 AAACAGCAGCTGAAAGAAAGGGGCAAGATTTCTGACTACATCTGATTAAGTTAC 360	
QY	361 ATTGTCAGATGAGGGGCGATGTTAGTACAGCAAGACCACTGAGAGACGTGAATTC 420	
DB	361 ATTGTCAGATGAGGGGCGATGTTAGTACAGCAAGACCACTGAGAGACGTGAATTC 420	
QY	421 GTGAGCCTTATTTCCGATATACAGAAATAGATCCCACTTTTCAAAAAGGGCATCA 480	
DB	421 GTGAGCCTTATTTCCGATATACAGAAATAGATCCCACTTTTCAAAAAGGGCATCA 480	
QY	481 GAGCTTGTAAAAGCAGTGGCGCTTGATACAAAGCAGAAATAAAGAGTGCATTTAAGA 540	
DB	481 GAGCTTGTAAAAGCAGTGGCGCTTGATACAAAGCAGAAATAAAGAGTGCATTTAAGA 540	
QY	541 GGCATCGAACAATGCGCAATTCGCAATTAAGCAATGATGTCATTAATTAACGGCAAG 600	
DB	541 GGCATCGAACAATGCGCAATTCGCAATTAAGCAATGATGTCATTAATTAACGGCAAG 600	
QY	601 CCTGATATAGGTGATGATGATGTTGGCGGTGAAATTTCAAAAGCGATGTGCTCAG 660	
DB	601 CCTGATATAGGTGATGATGATGTTGGCGGTGAAATTTCAAAAGCGATGTGCTCAG 660	
QY	661 AGCAGCTACGGGTTGATGAGCAAGGACGATCGTAGCGGTTCGCAATACAGGGCTTGAT 720	
DB	661 AGCAGCTACGGGTTGATGAGCAAGGACGATCGTAGCGGTTCGCAATACAGGGCTTGAT 720	
QY	721 ACAAGTCGCAATGACAGTTCGATGACATGAGGCTTCGCGGAGAAATTAAGTCATTAAT 780	
DB	721 ACAAGTCGCAATGACAGTTCGATGACATGAGGCTTCGCGGAGAAATTAAGTCATTAAT 780	
QY	781 GCATTTGAGACGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGAT 840	
DB	781 GCATTTGAGACGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGAT 840	
QY	841 TCCGATTTAGAAAGCGGCTCCATTAATAAAGAAATGCGGCTCAGGCGAATCTAGCTTC 900	
DB	841 TCCGATTTAGAAAGCGGCTCCATTAATAAAGAAATGCGGCTCAGGCGAATCTAGCTTC 900	

QY	901 CAATCTATCATGAGTACCGGTGGGGAAGCTTGGAGACTTCCGATCTGCAACCTTAA 960	
DB	901 CAATCTATCATGAGTACCGGTGGGGAAGCTTGGAGACTTCCGATCTGCAACCTTAA 960	
QY	961 TTACGCAAGCAATACAGTGTGCTGTCGCAAGATTCATCAAACTCCTGGGGAGACAGAG 1020	
DB	961 TTACGCAAGCAATACAGTGTGCTGTCGCAAGATTCATCAAACTCCTGGGGAGACAGAG 1020	
QY	1021 AATGGGGCTTACACAACAGATTCAGAAATGTGATGACTATGTGCGCAAAATGATATG 1080	
DB	1021 AATGGGGCTTACACAACAGATTCAGAAATGTGATGACTATGTGCGCAAAATGATATG 1080	
QY	1081 AGCATCTTTTGTGCTGCGGGAAATGAAGACCGAACCGAACCATGATGTCACAGGC 1140	
DB	1081 AGCATCTTTTGTGCTGCGGGAAATGAAGACCGAACCGAACCATGATGTCACAGGC 1140	
QY	1141 ACAGCTAAATAATGCAATTAACAGTGGAGCTACGCAAAACCTCCGCGCAAGCTTTGGGTCT 1200	
DB	1141 ACAGCTAAATAATGCAATTAACAGTGGAGCTACGCAAAACCTCCGCGCAAGCTTTGGGTCT 1200	
QY	1201 TATGGGACAAATATCAACATGTGGCACAGTTCTTTCAAGTGAACCGAACAGATGGA 1260	
DB	1201 TATGGGACAAATATCAACATGTGGCACAGTTCTTTCAAGTGAACCGAACAGATGGA 1260	
QY	1261 CGGATCAACCGGATGTGATGAGCACCGGGAAAGTTGATCATCTACGAAAGATCTTCTTT 1320	
DB	1261 CGGATCAACCGGATGTGATGAGCACCGGGAAAGTTGATCATCTACGAAAGATCTTCTTT 1320	
QY	1321 GCACCGGATTCCTCTCTGGGCGAAACATGACAGTAAATATGCAATAAGGTGGAACG 1380	
DB	1321 GCACCGGATTCCTCTCTGGGCGAAACATGACAGTAAATATGCAATAAGGTGGAACG 1380	
QY	1381 TCATGCTTACACCGATGTTGCTGGAAGACGTGCAACGCTTCTGAGCATTTTGTGAAA 1440	
DB	1381 TCATGCTTACACCGATGTTGCTGGAAGACGTGCAACGCTTCTGAGCATTTTGTGAAA 1440	
QY	1441 AACAGGCGATCAACCAAAAGCTTCTATTAATAAGGGGACATGATTCGGGCGAGCT 1500	
DB	1441 AACAGGCGATCAACCAAAAGCTTCTATTAATAAGGGGACATGATTCGGGCGAGCT 1500	
QY	1501 GACATCGGCTTGGCTAACCGGAACGGTAAACCAAGATGGGACAGTGAACATTTGGAATA 1560	
DB	1501 GACATCGGCTTGGCTAACCGGAACGGTAAACCAAGATGGGACAGTGAACATTTGGAATA 1560	
QY	1561 TCCTGAAAGTTGCTTATGTGAACGAGTCAAGTTCTTATCACACGCAAAAGCGAGCG 1620	
DB	1561 TCCTGAAAGTTGCTTATGTGAACGAGTCAAGTTCTTATCACACGCAAAAGCGAGCG 1620	
QY	1621 TACTGTTTACGCTACCGCGGCAAGCTTTGAAATCTCCCTGATATGCTGATGAGCC 1680	
DB	1621 TACTGTTTACGCTACCGCGGCAAGCTTTGAAATCTCCCTGATATGCTGATGAGCC 1680	
QY	1681 CCTGAGACCAACTGCTTCGTAACCGTTGTCATGATCTGACCTTGTCAATTAACGCT 1740	
DB	1681 CCTGAGACCAACTGCTTCGTAACCGTTGTCATGATCTGACCTTGTCAATTAACGCT 1740	
QY	1741 CCAATGCGACACAGTATGTAAGAAATGACTTTACTTCCGCAATCAATGATTAACGGGAT 1800	
DB	1741 CCAATGCGACACAGTATGTAAGAAATGACTTTACTTCCGCAATCAATGATTAACGGGAT 1800	
QY	1801 GGGCGCAATTAACGTAAGAAATGATTTATTAATGCAACCAAAAGCGGAGAGCTAATCAAT 1860	
DB	1801 GGGCGCAATTAACGTAAGAAATGATTTATTAATGCAACCAAAAGCGGAGAGCTAATCAAT 1860	
QY	1861 GAGGTACAGGCTTAAAGTACCGGTTGACCAAGACCTTCTGTTGGCAATTTGTAAT 1920	
DB	1861 GAGGTACAGGCTTAAAGTACCGGTTGACCAAGACCTTCTGTTGGCAATTTGTAAT 1920	
QY	1921 TAA 1923	
DB	1921 TAA 1923	

RESULT 2
LOCUS CQ888285 1923 bp DNA linear PAT 19-OCT-2004
DEFINITION Sequence 2 from Patent EP1466970.
ACCESSION CQ888285
VERSION CQ888285.1 GI:54304586
KEYWORDS
SOURCE Bacillus sp. KSM-KP43
ORGANISM Bacillus sp. KSM-KP43
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 Sato, T., Okuda, M., Koyama, S., Izawa, Y. and Kobayashi, T.
Alkaline protease
Patent: EP 1466970-A 2 13-OCT-2004;
Kao Corporation (JP)
FEATURES
source
1. 1923
/organism="Bacillus sp. KSM-KP43"
/mol_type="unassigned DNA"
/db_xref="taxon:109322"
1. 1923
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAH61612.1"
/db_xref="GI:54304587"
/translation="MRKKKVPISVLSAAIISTVALSNPAGANFDLDFKGIOTT
Y: TAKGSKQGGTGAAPLVSEENVKLPKGIQKLETPPANKLHIIQPNPILBETKO
OLEKTAKLIDYIPDYAIVIEGVGDKVASTIEHVESVPYPIYRIDPLETKAS
ELVKAVALDPTKQNKSEVOLARIGIOIAOPALSNVLYITAEKXWMDVARGVKKDV
AOSVGLYGGGQIVAVADPGLDGRBDSNHEA.PRGKITLALYALGTNNANDNGST
HVASGLKNSSTKGMAPOMLVFQSTIMSGGGLGGLPSLQTLFSDATSAKARHTN
SWGAIVAGATTTDSRVNDDVVRNDDMTILFAAGNBPNGGTISAAGTAKNAITVGATE
NMAPSGSYADININVAOPSSRGPTKXGR.IKPDVMA.PGFILISAPDSIFVGAITE
DSKYATMGSTNAPPIVAGVNAQLRHFVNRGITPEPSLLKALLAGAADIGLGP
GNGGMRVTLDEKLANVAVNSSLSTSOAKATSPATAGKPKILGIMVSDAPASTTA
SVTLVNDLIVITAPNGTQVYGNDFPSPNNDMDGRNENVEINAPQSGTTTIEVQA
YNPVGPQPTSLAIVN"
1. 618
619
sig_peptide
mat_peptide
ORIGIN
Query Match 100.0%; Score 1923; DB 6; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAAAAGAAAAGAGTGTGTTTATCTGTTTATCAGCTGACGAGATTTTGTGACT 60
DB 1 ATGAGAAAAGAAAAGAGTGTGTTTATCTGTTTATCAGCTGACGAGATTTTGTGACT 60
QY 61 GTTGCCTTAAGTAATCATCTGACGAGTGTGCAAGAAATTTGATCTGAATTCAAAGA 120
DB 61 GTTGCCTTAAGTAATCATCTGACGAGTGTGCAAGAAATTTGATCTGAATTCAAAGA 120
QY 121 ATTGACAGCAACATGCTTAAGGTTTCTCCAAAGCGGGGCACTGTGCTGCTGCT 180
DB 121 ATTGACAGCAACATGCTTAAGGTTTCTCCAAAGCGGGGCACTGTGCTGCTGCT 180
QY 181 TTTCTGTGGAATCTGAAAATGTGAAATCCCAAAAAGTTTGCAAGAAAGCTTGAACA 240
DB 181 TTTCTGTGGAATCTGAAAATGTGAAATCCCAAAAAGTTTGCAAGAAAGCTTGAACA 240
QY 241 GTTCCGGCAAAATATAAATCTCATATTATCAATTCATGACCAATTTTGAAGAAAC 300
DB 241 GTTCCGGCAAAATATAAATCTCATATTATCAATTCATGACCAATTTTGAAGAAAC 300
QY 301 AAACACAGCTGAAAAAACAAGGGGCAAAAGATTTCTGACATACCTGATTATGCTTAC 360
DB 301 AAACACAGCTGAAAAAACAAGGGGCAAAAGATTTCTGACATACCTGATTATGCTTAC 360
QY 361 ATTGTGAGATATGAGGGCGATGTTTAAGTCAGAACAGACCAATTGACGCTGATATCC 420
DB 361 ATTGTGAGATATGAGGGCGATGTTTAAGTCAGAACAGACCAATTGACGCTGATATCC 420

DB 361 ATTGTGAGATATGAGGGCGATGTTTAAGTCAGAACAGACCAATTGACGCTGATATCC 420
QY 421 GTTGAAGCTTATTTGCCGATATACAGAAATGATCCCAAGCTTTTACAAAAAGGGGCAATCA 480
DB 421 GTTGAAGCTTATTTGCCGATATACAGAAATGATCCCAAGCTTTTACAAAAAGGGGCAATCA 480
QY 481 GAGCTGTATAACAGAGGGCGCTTGAATCAAAACAGAAAAATAAGAGGGCAATTAGA 540
DB 481 GAGCTGTATAACAGAGGGCGCTTGAATCAAAACAGAAAAATAAGAGGGCAATTAGA 540
QY 541 GGCATCGAACAAATCGCAATTCGCAATAGCAATGATGCTATATATATACCGCAAG 600
DB 541 GGCATCGAACAAATTCGCAATTCGCAATAGCAATGATGCTATATATATACCGCAAG 600
QY 601 CTTGAGATTAAGTGATGATGATGATGTTGCCCGTGAATTTGTCAAGCGGATGTGGCTCAG 660
DB 601 CTTGAGATTAAGTGATGATGATGATGTTGCCCGTGAATTTGTCAAGCGGATGTGGCTCAG 660
QY 661 AGCAGCTACGGGTTGATATGACAAAGACAGATGCTAGCGGTTGCCGATACAGGGCTTGAT 720
DB 661 AGCAGCTACGGGTTGATATGACAAAGACAGATGCTAGCGGTTGCCGATACAGGGCTTGAT 720
QY 721 ACAAGTGCATATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACAAGTGCATATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCATTTGGACGACGAAATATGCGAATGATATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCATTTGGACGACGAAATATGCGAATGATATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGTTATTTGGAACCGGCTCCATTAATTAAGAAATGGCGCTCAGCGGAATCTAAGTCTTC 900
DB 841 TCCGTTATTTGGAACCGGCTCCATTAATTAAGAAATGGCGCTCAGCGGAATCTAAGTCTTC 900
QY 901 CAATCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 CAATCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 TTCAGCCAAAGCATACAGTGTGTCAGAAATTCATACAACTCTGGGGAGCAGCAGT 1020
DB 961 TTCAGCCAAAGCATACAGTGTGTCAGAAATTCATACAACTCTGGGGAGCAGCAGT 1020
QY 1021 AATGGGGCTTACCAACAATTCAGAAATGCGAATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATGGGGCTTACCAACAATTCAGAAATGCGAATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ACGATCTTTTGTGCTGCGCGGAAATGAAGACCGAACCGGCGAACATCACTGACACAGGC 1140
DB 1081 ACGATCTTTTGTGCTGCGCGGAAATGAAGACCGAACCGGCGAACATCACTGACACAGGC 1140
QY 1141 ACGCTTAATAATGCAATTAACATGCGAGCTACGGAACCTCCGCCCAAGCTTTGGGTC 1200
DB 1141 ACGCTTAATAATGCAATTAACATGCGAGCTACGGAACCTCCGCCCAAGCTTTGGGTC 1200
QY 1201 TATGCGGACAAATATCAACATGTGCAAGTCTCTTCAAGTGAACCGAACAAAGATGGA 1260
DB 1201 TATGCGGACAAATATCAACATGTGCAAGTCTCTTCAAGTGAACCGAACAAAGATGGA 1260
QY 1261 CGGATCAAAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 CGGATCAAAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 GCAACCGGATCTCTCTCTGCGGGAACCAATGACATGATTAATGATGATGATGATGATGATGAT 1380
DB 1321 GCAACCGGATCTCTCTCTGCGGGAACCAATGACATGATTAATGATGATGATGATGATGATGAT 1380
QY 1381 TCCATGCTACACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 TCCATGCTACACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1441 AACAGAGGATCAACCAAGCTTCTCTATTAATAAGCGGCACTGATGATGATGATGATGATGAT 1500
DB 1441 AACAGAGGATCAACCAAGCTTCTCTATTAATAAGCGGCACTGATGATGATGATGATGATGAT 1500

Qy	1501	GACATCGGCTTGGGCTACCCGACCGGTAA	CCAAAGATGGGAGCGAGTACATTGGATTAA	1560
Db	1501	GACATCGGCTTGGGCTACCCGACCGGTAA	CCAAAGATGGGAGCGAGTACATTGGATTAA	1560
Qy	1561	TCCCGTAAGCTTGCCCTATGTGTAACGAGTCCAGTTCTTATGACCAAGCCAAAAAGCGAGC	1620	
Db	1561	TCCCGTAAGCTTGCCCTATGTGTAACGAGTCCAGTTCTTATGACCAAGCCAAAAAGCGAGC	1620	
Qy	1621	TACTCGTTTACTGTACTAGTCCGGCAAGCCTTTGAAAATCTCCCTGGTATGGTCTGATGCC	1680	
Db	1621	TACTCGTTTACTGTACTAGTCCGGCAAGCCTTTGAAAATCTCCCTGGTATGGTCTGATGCC	1680	
Qy	1681	CCTGGGAGACCAACTGCTTCCGTAA	CAGTTGTCAATGATCTGAACTTGTCAATCCGCT	1740
Db	1681	CCTGGGAGACCAACTGCTTCCGTAA	CAGTTGTCAATGATCTGAACTTGTCAATCCGCT	1740
Qy	1741	CCAAATGGCACACAGTATGTAGGAAATGATCTTATCTTGCCATACATATGATTAATCTGGAT	1800	
Db	1741	CCAAATGGCACACAGTATGTAGGAAATGATCTTATCTTGCCATACATATGATTAATCTGGAT	1800	
Qy	1801	GGCCGCAATTAACGTAGAAATATGATTTATTTATATGACACCAACAAGCGGAGCGTATACATT	1860	
Db	1801	GGCCGCAATTAACGTAGAAATATGATTTATTTATATGACACCAACAAGCGGAGCGTATACATT	1860	
Qy	1861	GAGGTACAGGCTTATACGTACCGGTTGGACCAAGACCTTCTGCTTGGCAATTGTGAAT	1920	
Db	1861	GAGGTACAGGCTTATACGTACCGGTTGGACCAAGACCTTCTGCTTGGCAATTGTGAAT	1920	
Qy	1921	TAA	1923	
Db	1921	TAA	1923	
RESULT 3				
LOCUS	CO889309	1923 bp	DNA	linear
DEFINITION	Sequence 2 from Patent Epi1466562.			PAT 19-OCT-2004
ACCESSION	CO889309			
VERSION	CO889309.1	GI:54305176		
KEYWORDS				
SOURCE	Bacillus sp. KSM-KP43			
ORGANISM	Bacillus sp. KSM-KP43			
REFERENCE	1			
AUTHORS	Okuda,M., Sato,T., Takimura,Y., Sumitomo,N. and Kobayashi,T.			
TITLE	Alkaline protease			
JOURNAL	Patent: EP 1466962-A 2 13-OCT-2004;			
FEATURES	Kao Corporation (JP)			
SOURCE	location/Qualifiers			
	1..1923			
	/organism="Bacillus sp. KSM-KP43"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:109322"			
	1..1923			
	/note="unnamed protein product"			
	/codon_start=1			
	/transl_table=1			
	/protein_id="CAH1686.1"			
	/db_xref="GI:54305177"			
	1..1923			
	/translation="MRKKKVFVLVLSAAALISLVALSNPSAGANFPLDPFGIOIT			
	TDKAGSKOGOTGAALFLVESNVKLPGIKLKLTVPAANNLHLITQNPGLIETKO			
	QLEKTKAKITDIYIPDAYIVYVESGDVKSASTIEHVESPEPTPIYRIDPLFTKAS			
	ELYKAVALDTRKKRKEVOLAGIEQIOLFOFALISNDVLTITAAPEKVNNDVAKIIVADV			
	AQSSYGLYGGQQLVAADVTGLDTRNDSSMHEAFRKITVALVALGTINNANDTNGGT			
	HVAGSYGNGSTNGKGAPOANLVFOSIMSGGGLGGLPSNLQTLFQASAGARHITN			
	SNAPAVNGAYTTDSRVNDVYRKNDMTILEPAENEGNGGTSAPGTAKNAITVGATE			
	NLAPSPSYADININVAOFSSRGTRKDPGRKPDYVVAQPTTISAPSSIALPDSFPANH			
	DSYKAVMGCTSMATPIVAGNVADLREHPYKNRKITPKESILKRALTAGADTGLGPN			
	GGNGKGRVTLIDKSLNVAIVNNESSLSQCATYSPTATGKPKLISLVNSDAPSTTA			
	SVTLVNDLVLVITAPNGTGVGNDFTPSYNDMDGNNVENVFINAPDSGTIVTIVOA			
	LYNPVDPQPTSLAIIVN"			
	1..618			

sig_peptide

mat_peptide	619, .1920	/product="unnamed"
Query Match	100.0%	Score 1923; DB 6; Length 1923;
Best Local Similarity	100.0%	Pred. No. 0;
Matches 1923; Conservative	0;	Mismatches 0; Indels 0; Gaps 0
QY	1	ATGAGAAAGAGAAAAAGGTGTTTATCTGTTTATCAGCTGACGCATTTTGTGCAT
DB	1	ATGAGAAAGAGAAAAAGGTGTTTATCTGTTTATCAGCTGACGCATTTTGTGCAT
QY	61	GTTGCGTTAAGTAATTCATCTGACAGGTGTGCAGAGAAATTTGATCTGGATTTCAAGGA
DB	61	GTTGCGTTAAGTAATTCATCTGACAGGTGTGCAGAGAAATTTGATCTGGATTTCAAGGA
QY	121	ATTTCAGACAACTGAGTGGCTAAAGTTTCTCAAGCAGGGGCAAGCTGTGCTGCT
DB	121	ATTTCAGACAACTGAGTGGCTAAAGTTTCTCAAGCAGGGGCAAGCTGTGCTGCT
QY	181	TTTCTGTGTGAATCTGAAAATGTGAACTCCCAAAAGTTTTCAGAGAAAGCTTGAAACA
DB	181	TTTCTGTGTGAATCTGAAAATGTGAACTCCCAAAAGTTTTCAGAGAAAGCTTGAAACA
QY	241	GTTCCGGCAATTAATTAATCTCAATATTAATTCATTAATTCAGACCAATTTAGAGAAACA
DB	241	GTTCCGGCAATTAATTAATCTCAATATTAATTCATTAATTCAGACCAATTTAGAGAAACA
QY	301	AAACAGACGTGGAGAAAAACAGGGGCAAGAAATCTGCACTAATCTGATTAATGCTTAC
DB	301	AAACAGACGTGGAGAAAAACAGGGGCAAGAAATCTGCACTAATCTGATTAATGCTTAC
QY	361	ATTGTGAGTATGAGGGGCGATGTTTAAGTCAGACACAAACCAATTGAGACACGTGAAATCC
DB	361	ATTGTGAGTATGAGGGGCGATGTTTAAGTCAGACACAAACCAATTGAGACACGTGAAATCC
QY	421	GTTGAGCCTTATTTGCCGATATACAGAAATGATCCGACTTTTCAAAAAGGGCATCA
DB	421	GTTGAGCCTTATTTGCCGATATACAGAAATGATCCGACTTTTCAAAAAGGGCATCA
QY	481	GAGCTTGTAAGAAAGAGTGGCGCTTGATACAAAGCAGAAATTAAGAGGTGCATTAAGA
DB	481	GAGCTTGTAAGAAAGAGTGGCGCTTGATACAAAGCAGAAATTAAGAGGTGCATTAAGA
QY	541	GGCATCGAACAATATGCACAATTCGCAATTAAGCAATGATGCTAATATTACGGCAAG
DB	541	GGCATCGAACAATATGCACAATTCGCAATTAAGCAATGATGCTAATATTACGGCAAG
QY	601	CCTGAGTATTAAGTATGATGATGTGGCGGTGGAATTTGTCAAGCGGATGTGGCTCAG
DB	601	CCTGAGTATTAAGTATGATGATGTGGCGGTGGAATTTGTCAAGCGGATGTGGCTCAG
QY	661	AGCAGCTACGGGTTGTATGACAAAGACAGATCTGATGGTGTGCCGATACAGGGCTTGAT
DB	661	AGCAGCTACGGGTTGTATGACAAAGACAGATCTGATGGTGTGCCGATACAGGGCTTGAT
QY	721	ACAGGTGCAATGACAGTTGCATGATGATGACCTTCCGGGGGAAATTAATCTGCATTATAT
DB	721	ACAGGTGCAATGACAGTTGCATGATGATGACCTTCCGGGGGAAATTAATCTGCATTATAT
QY	781	GCATTGGAGCGAGCAATTAATGCAATGATATCGAATGGTCAATGGTACGCATGTGGCTGGC
DB	781	GCATTGGAGCGAGCAATTAATGCAATGATATCGAATGGTCAATGGTACGCATGTGGCTGGC
QY	841	TCGCTATTAAGAAACGGCTCCACTTAATTAAGAAATGGCGCTCAGGGCAATCTAGCTTTC
DB	841	TCGCTATTAAGAAACGGCTCCACTTAATTAAGAAATGGCGCTCAGGGCAATCTAGCTTTC
QY	901	CAATCTATCAATGAGTAAAGGGGTGGGGGACCTTGGAAGACATCTTGGAATCTGCAAACTTA
DB	901	CAATCTATCAATGAGTAAAGGGGTGGGGGACCTTGGAAGACATCTTGGAATCTGCAAACTTA
QY	961	TTGAGCCAGCATACAGTGTGTGTGCGCAAAATTCATCAAACTCTGGGGAGCAGAGT
DB	961	TTGAGCCAGCATACAGTGTGTGTGCGCAAAATTCATCAAACTCTGGGGAGCAGAGT

Db 961 TTCAGCCAGCATACAGTGTCTGTGCGCAAAATTCATACAAATCTCTGGGAGGAGCATGTG 1020
Qy 1021 AATGGGGCTTACACACAGATTCACAAAATGTGATGATCTATGTGGCGCAAAAATGATATG 1080
Db 1021 AATGGGGCTTACACACAGATTCACAAAATGTGATGATCTATGTGGCGCAAAAATGATATG 1080
Qy 1081 ACGATCTCTTTCGCTGCGGGGAAATGAAGGACCGGACGGGGGAAACCATCAGTGCACAGAGC 1140
Db 1081 ACGATCTCTTTCGCTGCGGGGAAATGAAGGACCGGACGGGGGAAACCATCAGTGCACAGAGC 1140
Qy 1141 ACGACTAAAATGCAATTAACAGTCGAGGCTACGAAAACTCCGCCCAAGCTTTGGGCTCT 1200
Db 1141 ACGACTAAAATGCAATTAACAGTCGAGGCTACGAAAACTCCGCCCAAGCTTTGGGCTCT 1200
Qy 1201 TATGGCGAATATCAACCATGTGGGACAGTTCTTTCAGTGGACCGCAAAAGATGGA 1260
Db 1201 TATGGCGAATATCAACCATGTGGGACAGTTCTTTCAGTGGACCGCAAAAGATGGA 1260
Qy 1261 CGGATCAAAACCGGATGTCAATGGGACCGGGGAAAGTTCAATACAGAAAGATCTTCTCT 1320
Db 1261 CGGATCAAAACCGGATGTCAATGGGACCGGGGAAAGTTCAATACAGAAAGATCTTCTCT 1320
Qy 1321 GCAACCGGATTCCTCTCTGCGGCGAAACCATGACAGTAATATGCATACATGGGTGAGCG 1380
Db 1321 GCAACCGGATTCCTCTCTGCGGCGAAACCATGACAGTAATATGCATACATGGGTGAGCG 1380
Qy 1381 TCCATGGCTACACCATTCGTTGCTGGAACGTTGCGACAGCTTCTGTAAGCAATTTTGAAA 1440
Db 1381 TCCATGGCTACACCATTCGTTGCTGGAACGTTGCGACAGCTTCTGTAAGCAATTTTGAAA 1440
Qy 1441 AACAGAGGATGACACCAAGCCTTCTATTTAAAGCGGACATGATGCGGTGACGCT 1500
Db 1441 AACAGAGGATGACACCAAGCCTTCTATTTAAAGCGGACATGATGCGGTGACGCT 1500
Qy 1501 GACATCGGCTTGGCTACCGGACGTAACCAAGATGGGAGAGATGACATTTGATTA 1560
Db 1501 GACATCGGCTTGGCTACCGGACGTAACCAAGATGGGAGAGATGACATTTGATTA 1560
Qy 1561 TCCCTGAACGTTGCTATGTGAACAGATCCAGTTCTCTATCCACAGCCCAAAAAGGACG 1620
Db 1561 TCCCTGAACGTTGCTATGTGAACAGATCCAGTTCTCTATCCACAGCCCAAAAAGGACG 1620
Qy 1621 TACTGTTTACTGCTACATGCGGCGCAAGCCTTGAATACTCCCTGATGATGCTGATGCG 1680
Db 1621 TACTGTTTACTGCTACATGCGGCGCAAGCCTTGAATACTCCCTGATGATGCTGATGCG 1680
Qy 1681 CCTGGGAGCAACATGCTTCCGTAACGCTTGTCAATGATCTGGACCTTGTCAATACCGCT 1740
Db 1681 CCTGGGAGCAACATGCTTCCGTAACGCTTGTCAATGATCTGGACCTTGTCAATACCGCT 1740
Qy 1741 CCAATATGCAACAGATATGAGAAATGAATTTACTTCCCATATCAATATTAATCTGGAT 1800
Db 1741 CCAATATGCAACAGATATGAGAAATGAATTTACTTCCCATATCAATATTAATCTGGAT 1800
Qy 1801 GGCCGCAATTAAGTGAATAATGATTTATTAATGCAACCAAGCGGAGCGATTAACAT 1860
Db 1801 GGCCGCAATTAAGTGAATAATGATTTATTAATGCAACCAAGCGGAGCGATTAACAT 1860
Qy 1861 GAGGTACAGGCTTATTAACGATCGGTTGACCAAGACCTTCTGTTGCAATTTGAT 1920
Db 1861 GAGGTACAGGCTTATTAACGATCGGTTGACCAAGACCTTCTGTTGCAATTTGAT 1920
Qy 1921 TAA 1923
Db 1921 TAA 1923

RESULT 4
AR368117 1923 bp DNA 1linear PAT 12-SEP-2003
LOCUS AR368117
DEFINITION Sequence 5 from patent US 6376227.
ACCESSION AR368117

VERSION AR368117.1 GI:34601778
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hltoml,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 5 23-APR-2002;
Kao Corporation; Tokyo;
JPX;
FEATURES
source Location/Qualifiers
1..1923
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1923; DB 6; Length 1923;
Best Local Similarity 100.0%; Pred.No.0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGAAAGAGAAAGAGGTTTATCTGTTTATCAGCTGACGCAATTTGTGACT 60
Db 1 ATGAGAAAGAGAAAGAGGTTTATCTGTTTATCAGCTGACGCAATTTGTGACT 60
Qy 61 GTTGCCTTAATGATATCATCTGAGGTGTGCAAGGAATTTGATCTGATTTCAAGGA 120
Db 61 GTTGCCTTAATGATATCATCTGAGGTGTGCAAGGAATTTGATCTGATTTCAAGGA 120
Qy 121 ATTGAGACAAACCTGATGCTAAAGTTCTCCAAACAGGGGACGCTGTGCTGCT 180
Db 121 ATTGAGACAAACCTGATGCTAAAGTTCTCCAAACAGGGGACGCTGTGCTGCT 180
Qy 181 TTTCTGCTGGAATCTGAAAATGTGAATCTCCAAAAGTTTGCAAGGAAGCTTGAACA 240
Db 181 TTTCTGCTGGAATCTGAAAATGTGAATCTCCAAAAGTTTGCAAGGAAGCTTGAACA 240
Qy 241 GTTCCGGCAATTAATCACTCCATATTATCAATGACCAATTTTGAAGAAACA 300
Db 241 GTTCCGGCAATTAATCACTCCATATTATCAATGACCAATTTTGAAGAAACA 300
Qy 301 AAACAGCAGCTGAAAAGAAACAGGGGCAAGATTTCTGACATACCTGATATGCTTAC 360
Db 301 AAACAGCAGCTGAAAAGAAACAGGGGCAAGATTTCTGACATACCTGATATGCTTAC 360
Qy 361 ATTGTGAGATGAGGGCGATGTTAAGTCAGCAACAGCAATTTGAGACGTGGAATCC 420
Db 361 ATTGTGAGATGAGGGCGATGTTAAGTCAGCAACAGCAATTTGAGACGTGGAATCC 420
Qy 421 GTGAGCCTTATTTGCGATATACAGATAGATCCCACTTTTCAAAAAGGGGATCA 480
Db 421 GTGAGCCTTATTTGCGATATACAGATAGATCCCACTTTTCAAAAAGGGGATCA 480
Qy 481 GAGCTTGTAAAGCAGTGGCGCTTGAATCAAGCGAAGAAATTAAGAGGTGCAATTAAG 540
Db 481 GAGCTTGTAAAGCAGTGGCGCTTGAATCAAGCGAAGAAATTAAGAGGTGCAATTAAG 540
Qy 541 GGCAATGAAACCAATGCAACATTTCCCAATTAAGATGATGCTATATTTACGGCAAG 600
Db 541 GGCAATGAAACCAATGCAACATTTCCCAATTAAGATGATGCTATATTTACGGCAAG 600
Qy 601 CCTGATATTAAGTGAATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAG 660
Db 601 CCTGATATTAAGTGAATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAG 660
Qy 661 ACGAGCTACGGGTGTGATGAGCAAGGACAGATGTAAGCGGTGCGATACAGGGCTGAT 720
Db 661 ACGAGCTACGGGTGTGATGAGCAAGGACAGATGTAAGCGGTGCGATACAGGGCTGAT 720
Qy 721 ACGAGTGCATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 ACGAGTGCATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 780

QY 781 GCATTGGGACGAGCAATTAATGCCAATGATACGAATGCTATGCTAAGCATGTGGCTGGC 840
DB 781 GCAATGGGACGAGCAATTAATGCAATGATACGAATGCTATGCTAAGCATGTGGCTGGC 840
QY 841 TCCGATTAGGAACCGGCTCCACTAATTAAGAAATGGCGCTCAGGGCAATCTAAGTCTTC 900
DB 841 TCCGATTAGGAACCGGCTCCACTAATTAAGAAATGGCGCTCAGGGCAATCTAAGTCTTC 900
QY 901 CAATCTATCATGATATAGCGGTGGGGGACTTGGAGGACTACCTTGAATCTGCAACCTTA 960
DB 901 CAATCTATCATGATATAGCGGTGGGGGACTTGGAGGACTACCTTGAATCTGCAACCTTA 960
QY 961 TTGAGCGAAGCATATAGTGTGCTGTCAGAAATTCATACAACTCTGGGGAGAGACAGT 1020
DB 961 TTGAGCGAAGCATATAGTGTGCTGTCAGAAATTCATACAACTCTGGGGAGAGACAGT 1020
QY 1021 AATGGGGCTTACACAAGATTCCAGAAATGTGATGACTATGTGGCAAAATGATATG 1080
DB 1021 AATGGGGCTTACACAAGATTCCAGAAATGTGATGACTATGTGGCAAAATGATATG 1080
QY 1081 AGGATCTTTTGGCTGCCGGAAATGAAGACCGAACGGGGAAACCATCATGTGCACCAAGC 1140
DB 1081 AGGATCTTTTGGCTGCCGGAAATGAAGACCGAACGGGGAAACCATCATGTGCACCAAGC 1140
QY 1141 ACAGCTAAATAATGCATTAACGTGGAGGCTAGGGAAACCTCCGGCCAGCTTTGGTCT 1200
DB 1141 ACAGCTAAATAATGCATTAACGTGGAGGCTAGGGAAACCTCCGGCCAGCTTTGGTCT 1200
QY 1201 TATGCGGACATATCAACCATGTGGCAAGTTCTTTCAGTGGACCGACAAAGATGGA 1260
DB 1201 TATGCGGACATATCAACCATGTGGCAAGTTCTTTCAGTGGACCGACAAAGATGGA 1260
QY 1261 CGGATTAACCGGATGTCTATGGCAACCGGAAAGTTCTATCTATCAGCAAGATCTTCTCT 1320
DB 1261 CGGATTAACCGGATGTCTATGGCAACCGGAAAGTTCTATCTATCAGCAAGATCTTCTCT 1320
QY 1321 GCAACCGGATCTCTCTCTGCGGCAACATGACAGTAAATATGATACATGCGTGAAG 1380
DB 1321 GCAACCGGATCTCTCTCTGCGGCAACATGACAGTAAATATGATACATGCGTGAAG 1380
QY 1381 TCCATGCGCTACACCGATCTGTGGAACGTGGCAAGCTTCTGATGCAATTTGTGAA 1440
DB 1381 TCCATGCGCTACACCGATCTGTGGAACGTGGCAAGCTTCTGATGCAATTTGTGAA 1440
QY 1441 AACAGAGGATCAACCAAGGCTTCTATTAAGCGGCACTGATTTGCCGTGCACT 1500
DB 1441 AACAGAGGATCAACCAAGGCTTCTATTAAGCGGCACTGATTTGCCGTGCACT 1500
QY 1501 GACATCGGCTTGGCTACCCGAAACGGTAAACCAAGATGGGAGAGATGCAATTTGATTA 1560
DB 1501 GACATCGGCTTGGCTACCCGAAACGGTAAACCAAGATGGGAGAGATGCAATTTGATTA 1560
QY 1561 TCCCTGAACGTTGCTATGTGAAGAGTCCAGTTCTTATCCACAGCCAAAGGCAAG 1620
DB 1561 TCCCTGAACGTTGCTATGTGAAGAGTCCAGTTCTTATCCACAGCCAAAGGCAAG 1620
QY 1621 TACTGTTTACTGCTACTGCGGCAAGGCTTGAATCTCCCTGATATGCTGATGCC 1680
DB 1621 TACTGTTTACTGCTACTGCGGCAAGGCTTGAATCTCCCTGATATGCTGATGCC 1680
QY 1681 CTGTGAGACACACTGCTCCGTAACGCTTGTCAATGATCTGAGCCTTGTATCCGCT 1740
DB 1681 CTGTGAGACACACTGCTCCGTAACGCTTGTCAATGATCTGAGCCTTGTATCCGCT 1740
QY 1741 CCAATGCGACACAGATATGTAAGAAATGACTTTATCTTGCATATCAATGATTAAGTAT 1800
DB 1741 CCAATGCGACACAGATATGTAAGAAATGACTTTATCTTGCATATCAATGATTAAGTAT 1800
QY 1801 GGGCGCAATTAAGTAAATATGATTTATTAATGACACAAAGGGGAGCTATTAATTA 1860
DB 1801 GGGCGCAATTAAGTAAATATGATTTATTAATGACACAAAGGGGAGCTATTAATTA 1860

QY 1861 GAGTACAGGCTTATTAACGTACCGGTTGGACACAGACCTTCTCGTTGGCAATTTGAT 1920
DB 1861 GAGTACAGGCTTATTAACGTACCGGTTGGACACAGACCTTCTCGTTGGCAATTTGAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923
RESULT 5
ARS62465
LOCUS ARS62465 1923 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 5 from patent US 6759228.
ACCESSION ARS62465
VERSION ARS62465.1 GI:53976506
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hileomi,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6759228-A 5 06-JUL-2004;
Kao Corporation; Tokyo;
JFK;
FEATURES
source location/Qualifiers
1..1923
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1923; DB 6; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGACGATTTGTGCACT 60
DB 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGACGATTTGTGCACT 60
QY 61 GTTCCGTTAAGTAATCCATCTGACAGTGTGCAAGAAATTTTGAATCTGCAATTTCAAGGA 120
DB 61 GTTCCGTTAAGTAATCCATCTGACAGTGTGCAAGAAATTTTGAATCTGCAATTTCAAGGA 120
QY 121 ATTCAACAAACACTGATGCTAAAGTTTCTCCAAAGAGGGGAGACTGTGCTGCT 180
DB 121 ATTCAACAAACACTGATGCTAAAGTTTCTCCAAAGAGGGGAGACTGTGCTGCT 180
QY 181 TTTCTGATGGAATCTGAAATGTGAACCTCCAAAGGTTTGAAGAAAGCTTGAACA 240
DB 181 TTTCTGATGGAATCTGAAATGTGAACCTCCAAAGGTTTGAAGAAAGCTTGAACA 240
QY 241 GTCCCGCAATTAATAACTCAATATATCAATTCATATGACCAATTTTGAAGAAACA 300
DB 241 GTCCCGCAATTAATAACTCAATATATCAATTCATATGACCAATTTTGAAGAAACA 300
QY 241 GTCCCGCAATTAATAACTCAATATATCAATTCATATGACCAATTTTGAAGAAACA 300
DB 241 GTCCCGCAATTAATAACTCAATATATCAATTCATATGACCAATTTTGAAGAAACA 300
QY 301 AAACAGAGCTGGAAGAAACAGGGGCAAAAGTTCTGCACTTACCTGATTAAGCTTAC 360
DB 301 AAACAGAGCTGGAAGAAACAGGGGCAAAAGTTCTGCACTTACCTGATTAAGCTTAC 360
QY 361 ATTGTGAGATGAGGGGCAATGTTAAGTCAAGAACAGCACTTGGAGCGTGAATCC 420
DB 361 ATTGTGAGATGAGGGGCAATGTTAAGTCAAGAACAGCACTTGGAGCGTGAATCC 420
QY 421 GTGAGGCTTAATTTGCCGATATACAGAAATGATCCCAAGCTTTTCAAAAGGGGATCA 480
DB 421 GTGAGGCTTAATTTGCCGATATACAGAAATGATCCCAAGCTTTTCAAAAGGGGATCA 480
QY 481 GAGCTTGAATAACAGATGGGGCTTGAATACAAAGCAAAATTAAGAAGGTGCAATTAAGA 540
DB 481 GAGCTTGAATAACAGATGGGGCTTGAATACAAAGCAAAATTAAGAAGGTGCAATTAAGA 540
QY 541 GGCATGAACAATTCGACAAATTCGCAATTAAGCAATGATGTCTATATTAAGCGAAG 600

```

Db      541 GGCATGCAAAATCGCAATTCGCAATAGCAATATATGTCATATATATTCGCGCAAG 600
Qy      601 CCGAGATATAGAGTGAATGATGTTGCCGTGGAAATTCMAAGCGATGTGGCTCAG 660
Db      601 CCGAGATATAGAGTGAATGATGTTGCCGTGGAAATTCMAAGCGATGTGGCTCAG 660
Qy      661 AGCAATCAAGGATGTATGCAAGGACAGATGCTAGCGGTGGCGATACAGGCTTGAAT 720
Db      661 AGCAATCAAGGATGTATGCAAGGACAGATGCTAGCGGTGGCGATACAGGCTTGAAT 720
Qy      721 ACAGGTGCAATGACAGTTGATGATGAAGCCTTCGCGGAAATTTACTGCAATATAT 780
Db      721 ACAGGTGCAATGACAGTTGATGATGAAGCCTTCGCGGAAATTTACTGCAATATAT 780
Qy      781 GCATTTGGGACGACGAATTAATGCCAATGATACGAATGTCAATGTCGCAATGTGGCTGC 840
Db      781 GCATTTGGGACGACGAATTAATGCCAATGATACGAATGTCAATGTCGCAATGTGGCTGC 840
Qy      841 TCGGTATTAGAAACGGCTCCACTAATAAGAAATGGCGCTCAGCGGAAATCTAGTCTTC 900
Db      841 TCGGTATTAGAAACGGCTCCACTAATAAGAAATGGCGCTCAGCGGAAATCTAGTCTTC 900
Qy      901 CAATCTATCAATGATAGCGGTGGGGAATTCGAGGACTACCTTCGAAATCTGCAAACTTTA 960
Db      901 CAATCTATCAATGATAGCGGTGGGGAATTCGAGGACTACCTTCGAAATCTGCAAACTTTA 960
Qy      961 TTCAGCCAGACATACAGTGTGTCGCAATTAATCAAACTCTGCGGAGACAGCATGT 1020
Db      961 TTCAGCCAGACATACAGTGTGTCGCAATTAATCAAACTCTGCGGAGACAGCATGT 1020
Qy      1021 AATGGGGCTTACACACAGTTCCAGAAATGTGATGATCTATGTGGCCAAAATGATATG 1080
Db      1021 AATGGGGCTTACACACAGTTCCAGAAATGTGATGATCTATGTGGCCAAAATGATATG 1080
Qy      1081 AAGTGGGCTTACACACAGTTCCAGAAATGTGATGATCTATGTGGCCAAAATGATATG 1080
Db      1081 AAGTGGGCTTACACACAGTTCCAGAAATGTGATGATCTATGTGGCCAAAATGATATG 1080
Qy      1081 ACGATCTTTTCCGTCCGCGGAAATGAAAGACCGAACCGGAAACCATCACTGACACAGGC 1140
Db      1081 ACGATCTTTTCCGTCCGCGGAAATGAAAGACCGAACCGGAAACCATCACTGACACAGGC 1140
Qy      1141 ACGCTTAAATCAATTAACAGTGTGACGAAACCTCGCGCCAACTTTGGGCTCT 1200
Db      1141 ACGCTTAAATCAATTAACAGTGTGACGAAACCTCGCGCCAACTTTGGGCTCT 1200
Qy      1201 TATGCGGACAAATCAACATGTGGGACAGTTCTCTCAAGTGAACCGAACAGATATGA 1260
Db      1201 TATGCGGACAAATCAACATGTGGGACAGTTCTCTCAAGTGAACCGAACAGATATGA 1260
Qy      1261 CGGATCAAAACCGATGTCATGCGACCGGGAACGTTCACTACAGCAAGATCTTCTCTT 1320
Db      1261 CGGATCAAAACCGATGTCATGCGACCGGGAACGTTCACTACAGCAAGATCTTCTCTT 1320
Qy      1321 GCAACCGGATCTCTCTTCTGGGCGAACCATGACAGTAATATATCATATGAGTGAACG 1380
Db      1321 GCAACCGGATCTCTCTTCTGGGCGAACCATGACAGTAATATATCATATGAGTGAACG 1380
Qy      1381 TCCATGCGCTACACCGATCTGTTGCTGGAAAGTGGCACAGTTTGTGAGCAATTTTGAAA 1440
Db      1381 TCCATGCGCTACACCGATCTGTTGCTGGAAAGTGGCACAGTTTGTGAGCAATTTTGAAA 1440
Qy      1381 TCCATGCGCTACACCGATCTGTTGCTGGAAAGTGGCACAGTTTGTGAGCAATTTTGAAA 1440
Db      1381 TCCATGCGCTACACCGATCTGTTGCTGGAAAGTGGCACAGTTTGTGAGCAATTTTGAAA 1440
Qy      1441 AACAGAGGCTACACCAAGCTTCTATTAATAAGCGGCACTGATGCGGTGACGCT 1500
Db      1441 AACAGAGGCTACACCAAGCTTCTATTAATAAGCGGCACTGATGCGGTGACGCT 1500
Qy      1501 GACATGCGGCTTGGCTACCGGAACGGTAAACCAAGATGGGGAAGAGTGAATTTGAA 1560
Db      1501 GACATGCGGCTTGGCTACCGGAACGGTAAACCAAGATGGGGAAGAGTGAATTTGAA 1560
Qy      1561 TCCCTGAAGGTTGCTATGTAAGAGTCAAGTTCTCTATCCACAGCCAAAAGGAGCG 1620
Db      1561 TCCCTGAAGGTTGCTATGTAAGAGTCAAGTTCTCTATCCACAGCCAAAAGGAGCG 1620
Qy      1621 TACTCGTTTACTGCTACGCGGCAAGCCTTTGAAAATCTCCCTGATATGTGATGCG 1680
Db      1621 TACTCGTTTACTGCTACGCGGCAAGCCTTTGAAAATCTCCCTGATATGTGATGCG 1680

```

```

Db      1621 TACTCGTTTACTGCTACGCGGCAAGCCTTTGAAAATCTCCCTGATATGTGATGCG 1680
Qy      1681 CCGCGACGACAACTGCTCCGTAACGCTTGTCAATGATGTGACCTTTCATTAACGCT 1740
Db      1681 CCGCGACGACAACTGCTCCGTAACGCTTGTCAATGATGTGACCTTTCATTAACGCT 1740
Qy      1741 CCAATGGGACACAGATATGAGAAATGAATTTACTTCCGCAATCAATGATTAATGAGAT 1800
Db      1741 CCAATGGGACACAGATATGAGAAATGAATTTACTTCCGCAATCAATGATTAATGAGAT 1800
Qy      1801 GCGCGAATTAACCTAGAAATATGATTTATTAATGACACCAAAAGCGGACGTATACAT 1860
Db      1801 GCGCGAATTAACCTAGAAATATGATTTATTAATGACACCAAAAGCGGACGTATACAT 1860
Qy      1861 GAGGTACAGGCTTATTAACCTACCGGTGGACCAACAGACCTTCTCGTTGGCAATTTGAA 1920
Db      1861 GAGGTACAGGCTTATTAACCTACCGGTGGACCAACAGACCTTCTCGTTGGCAATTTGAA 1920
Qy      1921 TAA 1923
Db      1921 TAA 1923

RESULT 6
AR368118
LOCUS AR368118 1923 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 7 from patent US 6376227.
ACCESSION AR368118
VERSION AR368118.1 GI:34601779
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TAKAIWA,M., OKUDA,M., SAEKI,K., KUBOTA,H., HITOMI,J., KAGEYAMA,Y.,
SHIKATA,S. and NOMURA,M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 7 23-APR-2002;
Kao Corporation; Tokyo,
JPX;
FEATURES
source
location/Qualifiers
1..1923
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 99.8%; Score 1919.8; DB 6; Length 1923;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGCAAGCAATTTGTCGACT 60
Db 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGCAAGCAATTTGTCGACT 60
Qy 61 GTTGGCTTAAGTATCATCTGACAGGTGTGCAAGAAATTTGATGTGAAATTTCAAGGA 120
Db 61 GTTGGCTTAAGTATCATCTGACAGGTGTGCAAGAAATTTGATGTGAAATTTCAAGGA 120
Qy 121 ATTGAGACAAACAGATGCTAAAGTTTCTCAAGCAGGGGACAGCTGCTGCTGCT 180
Db 121 ATTGAGACAAACAGATGCTAAAGTTTCTCAAGCAGGGGACAGCTGCTGCTGCTGCT 180
Qy 181 TTTCTGCTGGAATCTGAAATATGAAACTCCCAAAAGGTTTGCAGAAAGAGCTTGAACA 240
Db 181 TTTCTGCTGGAATCTGAAATATGAAACTCCCAAAAGGTTTGCAGAAAGAGCTTGAACA 240
Qy 241 GTCCCGGCAATTAATTAATCTCCATTTATTCATTAATGACCAATTTTGAAGAAACA 300
Db 241 GTCCCGGCAATTAATTAATCTCCATTTATTCATTAATGACCAATTTTGAAGAAACA 300
Qy 301 AAAACAGCAGCTGGAAGAAACAGGGGCAAGATTTCTGACTACATACCTGATTATGCTTAC 360
Db 301 AAAACAGCAGCTGGAAGAAACAGGGGCAAGATTTCTGACTACATACCTGATTATGCTTAC 360

```

QY 361 ATTGTCAGTATGAGGGCGATGTTAAGTCAGCAACAGGACCATGAGACGTGGAATCC 420
DB 361 ATTGTCAGTATGAGGGCGATGTTAAGTCAGCAACAGGACCATGAGACGTGGAATCC 420
QY 421 GTGAGACCTTTATTTGCGATATACAGAAATAGATCCCAGCTTTTCACAAAAGGGGCAATCA 480
DB 421 GTGAGACCTTTATTTGCGATATACAGAAATAGATCCCAGCTTTTCACAAAAGGGGCAATCA 480
QY 481 GAGCTGTAAAGCAATGCGCTTGATACAAAGCAGAAAAATTAAGAGGTGCATTTAA 540
DB 481 GAGCTGTAAAGCAATGCGCTTGATACAAAGCAGAAAAATTAAGAGGTGCATTTAA 540
QY 541 GGCATGCAACAAATCGCACAATTCCGATTAAGCATATGTGCTATATATTAACGGCAAG 600
DB 541 GGCATGCAACAAATCGCACAATTCCGATTAAGCATATGTGCTATATATTAACGGCAAG 600
QY 601 CCTGAGTATAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 CCTGAGTATAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 AGCAGTACAGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 AGCAGTACAGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 ACAGGTGCAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACAGGTGCAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCATTGGGACGACGCAATATATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCATTGGGACGACGCAATATATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGATATTAGAAACGGCTCCACTAATTAAGGAATGGGCGCTCAGGCGAATCTACTCTTC 900
DB 841 TCCGATATTAGAAACGGCTCCACTAATTAAGGAATGGGCGCTCAGGCGAATCTACTCTTC 900
QY 901 CAATCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 CAATCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 TTCAACGCAAGCATACATGCTGCTGCGCAGAAATTCATACAACTCTCGGGGACAGCAGTG 1020
DB 961 TTCAACGCAAGCATACATGCTGCTGCGCAGAAATTCATACAACTCTCGGGGACAGCAGTG 1020
QY 1021 AATGGGGCTTACACAAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATGGGGCTTACACAAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ACGATCCTTTTCTGCGCGGGAATGAGGACCGGACCGGACCGGACCGGACCGGACCGGAC 1140
DB 1081 ACGATCCTTTTCTGCGCGGGAATGAGGACCGGACCGGACCGGACCGGACCGGACCGGAC 1140
QY 1141 ACAAGTAAATGCAATACATGCTGCGAGCTACGGAACCTTCGCGCAAGCTTTGGGCT 1200
DB 1141 ACAAGTAAATGCAATACATGCTGCGAGCTACGGAACCTTCGCGCAAGCTTTGGGCT 1200
QY 1201 TATGCGGACCAATATACATGCTGCGAGCTTCTTCAAGTGGACCGGACCAAGGATGGA 1260
DB 1201 TATGCGGACCAATATACATGCTGCGAGCTTCTTCAAGTGGACCGGACCAAGGATGGA 1260
QY 1261 CGGATTAACCGGATGCTATGCGACCGGGAAGCTTCACTACTATACGAAAGATCTTCTCT 1320
DB 1261 CGGATTAACCGGATGCTATGCGACCGGGAAGCTTCACTACTATACGAAAGATCTTCTCT 1320
QY 1321 GCACCGGATTCCTCTTCTGCGGCAACCATGACAGTAAATATGATATATGATGATGATGAT 1380
DB 1321 GCACCGGATTCCTCTTCTGCGGCAACCATGACAGTAAATATGATATATGATGATGATGAT 1380
QY 1381 TCCATGCGTACACCGATGCTGCTGGAACGTCGACAGCTTCGAGACATTTTGTGAAA 1440
DB 1381 TCCATGCGTACACCGATGCTGCTGGAACGTCGACAGCTTCGAGACATTTTGTGAAA 1440

QY 1441 AACAGAGCATCACCAAGGCTTCTCTATTTAAAGGCGACATGATCCGGCTGACCT 1500
DB 1441 AACAGAGCATCACCAAGGCTTCTCTATTTAAAGGCGACATGATCCGGCTGACCT 1500
QY 1501 GACATCGGCTTTGGCTACCCGGAACGGTAAACCAAGGATGGGACGAGTACATTTGATAAA 1560
DB 1501 GACATCGGCTTTGGCTACCCGGAACGGTAAACCAAGGATGGGACGAGTACATTTGATAAA 1560
QY 1561 TCCCTGAACGTTGCTTATGATGACAGTTCAGTCTTATATCAACGCAAAAGCGACG 1620
DB 1561 TCCCTGAACGTTGCTTATGATGACAGTTCAGTCTTATATCAACGCAAAAGCGACG 1620
QY 1621 TACTGTTTACGCTACTACCGGCAAGGCTTTGAAATCTCCCTGATATGCTGATGATG 1680
DB 1621 TACTGTTTACGCTACTACCGGCAAGGCTTTGAAATCTCCCTGATATGCTGATGATG 1680
QY 1681 CCTGCGACCAACCTGCTTCGATACAGCTTGTCAATGATCTGACCTTGTATTAACGCT 1740
DB 1681 CCTGCGACCAACCTGCTTCGATACAGCTTGTCAATGATCTGACCTTGTATTAACGCT 1740
QY 1741 CCAATGCGACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 1741 CCAATGCGACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 1801 GACCGCAATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1801 GACCGCAATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 1861 GAGGTACAGGCTTATTAACGATCCGCTTGACACACAGACCTTCTGCTGGCAATTTG 1920
DB 1861 GAGGTACAGGCTTATTAACGATCCGCTTGACACACAGACCTTCTGCTGGCAATTTG 1920
QY 1921 TAA 1923
DB 1921 TAA 1923

RESULT 7
ARS62466
LOCUS ARS62466 1923 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 7 from patent US 6759228.
ACCESSION ARS62466
VERSION ARS62466.1 GI:53976507
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hltomi,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNALS Patent: US 6759228-A 7 06-JUL-2004;
Kao Corporation; Tokyo;
JPK;

FEATURES
source location/Qualifiers
1..1923
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 99.8%; Score 1919.8; DB 6; Length 1923;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCACTGACGAGATTTTGTGACT 60
DB 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCACTGACGAGATTTTGTGACT 60
QY 61 GTTGCGTTAAGTATCATCATGAGGTGCGCAAGAAATTTGATCTGATTTCAAGGA 120
DB 61 GTTGCGTTAAGTATCATCATGAGGTGCGCAAGAAATTTGATCTGATTTCAAGGA 120
QY 121 ATTACAGCAACACTGATGCTAAAGTTTCTCCAGACAGGGGACAGACTGCTGCTGCT 180

Db 121 ATTGAGCAACAACTGATGCTTAAAGTTTCTCCAGACGGGCGAGACTGGTCTGCTGCT 180
 Qy 181 TTTCTGTGTGAATCTGAAAAATGTGAAACTCCCAAAAGTTTGCAGAAAGATTGAAACA 240
 Db 181 TTTCTGTGTGAATCTGAAAAATGTGAAACTCCCAAAAGTTTGCAGAAAGATTGAAACA 240
 Qy 241 GTCCCGGCAAAATTAATAACTCCATTTATTCAGATTGCAACCAATTTTGAAGAAACA 300
 Db 241 GTCCCGGCAAAATTAATAACTCCATTTATTCAGATTGCAACCAATTTTGAAGAAACA 300
 Qy 301 AAACAGACGCTGAAAAAAACAGGGGCAAAAGATTCTCGACTACATACCTGATTAATGCTTAC 360
 Db 301 AAACAGACGCTGAAAAAAACAGGGGCAAAAGATTCTCGACTACATACCTGATTAATGCTTAC 360
 Qy 361 ATTGTGAGATGAGAGGCGCATGTTAAGTCAGCAACAGCAACAATTGAGCAGTGAATCC 420
 Db 361 ATTGTGAGATGAGAGGCGCATGTTAAGTCAGCAACAGCAACAATTGAGCAGTGAATCC 420
 Qy 421 GTGGAACCTTATTTTGGCGAATATACAGATAGATCCCGAGCTTTTCAAAAAGGGGCAATCA 480
 Db 421 GTGGAACCTTATTTTGGCGAATATACAGATAGATCCCGAGCTTTTCAAAAAGGGGCAATCA 480
 Qy 481 GAGCTGTAAAAACAGTGGCGCTTGAATACAAAGCAAAAAATTAAGAGTGAATTAAGA 540
 Db 481 GAGCTGTAAAAACAGTGGCGCTTGAATACAAAGCAAAAAATTAAGAGTGAATTAAGA 540
 Qy 541 GGCATCGAAACAAATCGCAACATTCGCAATTAAGCAATGATGTGCTATATTAACCGCAAG 600
 Db 541 GGCATCGAAACAAATCGCAACATTCGCAATTAAGCAATGATGTGCTATATTAACCGCAAG 600
 Qy 601 CCTGATATTAAGTGTATGATATGATGTTGGCGGTGGAATTTGTCAAAGCGGATGGCTCAG 660
 Db 601 CCTGATATTAAGTGTATGATATGATGTTGGCGGTGGAATTTGTCAAAGCGGATGGCTCAG 660
 Qy 661 AGCAGCTACGGGTTGTATGAGCAAGGACAGATGCTGACGGGTGCGATACAGGGCTTGAAT 720
 Db 661 AGCAGCTACGGGTTGTATGAGCAAGGACAGATGCTGACGGGTGCGATACAGGGCTTGAAT 720
 Qy 721 ACAGGTGCGCAATGACAGTTGATGATGATGAAGCCTTCGCGGGAAAAATTAATCTGATTAAT 780
 Db 721 ACAGGTGCGCAATGACAGTTGATGATGATGAAGCCTTCGCGGGAAAAATTAATCTGATTAAT 780
 Qy 781 GCATTTGGGACGCAATTAATGCAATGATAGCAATGATGATGATGATGATGATGATGATGAT 840
 Db 781 GCATTTGGGACGCAATTAATGCAATGATAGCAATGATGATGATGATGATGATGATGATGAT 840
 Qy 841 TCCGATTTAGGAAACGGGCTCCACTAATTAAGAAATGGCGCTCAGGGGCAATCTAAGTCTTC 900
 Db 841 TCCGATTTAGGAAACGGGCTCCACTAATTAAGAAATGGCGCTCAGGGGCAATCTAAGTCTTC 900
 Qy 901 CAATCTATCATGATAGCGGTGGGGGACTTGGAGACTACCTTCGATCTGCAAACTTTA 960
 Db 901 CAATCTATCATGATAGCGGTGGGGGACTTGGAGACTACCTTCGATCTGCAAACTTTA 960
 Qy 961 TTTCAGCGCAAGCATACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 Db 961 TTTCAGCGCAAGCATACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 Qy 1021 AATGGGGCTTACACACAGATTCGCAAAATGAGGATGATGATGATGATGATGATGATGATGAT 1080
 Db 1021 AATGGGGCTTACACACAGATTCGCAAAATGAGGATGATGATGATGATGATGATGATGATGAT 1080
 Qy 1081 ACAGATCTTTTCCGCTCCGGGATGAGAGCCGAGCGGGAACCATCAGTGCACAGGCG 1140
 Db 1081 ACAGATCTTTTCCGCTCCGGGATGAGAGCCGAGCGGGAACCATCAGTGCACAGGCG 1140
 Qy 1141 ACAGCTAAAAATGCAATTAACAGTGCAGGCTACGGAACCTCCGCCCAAGCTTTGGGTCT 1200
 Db 1141 ACAGCTAAAAATGCAATTAACAGTGCAGGCTACGGAACCTCCGCCCAAGCTTTGGGTCT 1200
 Qy 1201 TATGGGCAATTAATGAACATGAGGCAAGTTCTCTTCAAGTGAACGCAAGAGATGGA 1260
 Db 1201 TATGGGCAATTAATGAACATGAGGCAAGTTCTCTTCAAGTGAACGCAAGAGATGGA 1260

Db 1201 TATGGGCAATTAATGAACATGAGGCAAGTTCTCTTCAAGTGAACGCAAGAGATGGA 1260
 Qy 1261 CGGATCAAAACCGGATGATGAGCAACGGGAACTTCAATCATACAGCAAGATCTTCTCT 1320
 Db 1261 CGGATCAAAACCGGATGATGAGCAACGGGAACTTCAATCATACAGCAAGATCTTCTCT 1320
 Qy 1321 GCAACGGATTTCCCTCTTGGGGGACATGACAGTAATTAATGATGATGATGATGATGATGAT 1380
 Db 1321 GCAACGGATTTCCCTCTTGGGGGACATGACAGTAATTAATGATGATGATGATGATGATGAT 1380
 Qy 1381 TCCATGGCTACACCGATGCTGTTGGGAAAGTGGCAACGCTTGTGAGCATTTTGTGAAA 1440
 Db 1381 TCCATGGCTACACCGATGCTGTTGGGAAAGTGGCAACGCTTGTGAGCATTTTGTGAAA 1440
 Qy 1441 AACAGAGGATCAACCAAGCTTCTGATTAATAAGCGGCACTGATTTGCCGTGACGCT 1500
 Db 1441 AACAGAGGATCAACCAAGCTTCTGATTAATAAGCGGCACTGATTTGCCGTGACGCT 1500
 Qy 1501 GACATGGGCTTGGCTACCCGAAACGGTAACCAAGATGGGGACAGTGAATGATGATGAT 1560
 Db 1501 GACATGGGCTTGGCTACCCGAAACGGTAACCAAGATGGGGACAGTGAATGATGATGAT 1560
 Qy 1561 TCCCTGAACGTTGCTATGTAACGAGTCAAGTCTCTATTCACAGCCCAAAAAGGACG 1620
 Db 1561 TCCCTGAACGTTGCTATGTAACGAGTCAAGTCTCTATTCACAGCCCAAAAAGGACG 1620
 Qy 1621 TACTGTTTACTGCTATGCTACGCGGCAAGCTTTGAAAAATCTCCGTGATGCTGATGCC 1680
 Db 1621 TACTGTTTACTGCTATGCTACGCGGCAAGCTTTGAAAAATCTCCGTGATGCTGATGCC 1680
 Qy 1681 CCTGCGAGCAACATGCTCTCCGTAAAGCTTTGCAATGATGATGATGATGATGATGATGAT 1740
 Db 1681 CCTGCGAGCAACATGCTCTCCGTAAAGCTTTGCAATGATGATGATGATGATGATGATGAT 1740
 Qy 1741 CCAATGAGCAACAGTATGAGAAATGACTTATCTTGCSCATPAACAATGATTAACGAGAT 1800
 Db 1741 CCAATGAGCAACAGTATGAGAAATGACTTATCTTGCSCATPAACAATGATTAACGAGAT 1800
 Qy 1801 GGCAGCAATTAACGTAGAAAAATGATTAATTAATGACCAACAAAGCGGACGATACAT 1860
 Db 1801 GGCAGCAATTAACGTAGAAAAATGATTAATTAATTAATGACCAACAAAGCGGACGATACAT 1860
 Qy 1861 GAGGTACAGCTTATTAACGTACCGGTGAGCAACAGCTTCTCGTGGCAATTTGAT 1920
 Db 1861 GAGGTACAGCTTATTAACGTACCGGTGAGCAACAGCTTCTCGTGGCAATTTGAT 1920
 Qy 1921 TAA 1923
 Db 1921 TAA 1923

RESULT 8
 AB084155 1923 bp DNA linear BCT 02-SEP-2003
 LOCUS Bacillus sp. KSM-9865 protease gene complete cds.
 DEFINITION AB084155
 ACCESSION AB084155.1 GI:34392386
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Bacillus sp. KSM-9865
 Bacillus sp. KSM-9865
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE
 1 Okuda, M., Saeki, K. and Kobayashi, T.
 TITLE Bacillus sp. KSM-9865 protease gene
 JOURNAL Published Only in Database (2003)
 REFERENCE
 2 (bases 1 to 1923)
 AUTHORS Okuda, M., Saeki, K. and Kobayashi, T.
 TITLE Direct Submision
 JOURNAL Submitted (18-Apr-2002) Mitsuyoshi Okuda, Kao corporation,
 Biological Science Laboratory, Ichikaimachi Akabane 2606, Haga,
 Tochigi 321-3497, Japan (E-mail: okuda.mitsuyoshi@kao.co.jp,
 Tel: 81-285-68-7543, Fax: 81-285-68-7547)
 FEATURES
 Location/Qualifiers

source 1. .1923
/organism="Bacillus sp. KSM-9865"
/mol_type="genomic DNA"
/strain="KSM-9865"
/db_xref="taxon:192495"
1. .1923
/codon_start=1
/transl_table=11
/product="protease"
/protein_id="BAC82522.1"
/db_xref="GI:34392387"
/translation="MRKKKKVPLSVLSAAAILSTVALSNPSAGANPDLDPKGIQTI
TDAKPSKQSGTGAALFLVESBNVKKPKLQKLETPVANNKHLIVQFNPILSTKQ
OLEKTKALIDYIPDYAVIVEYGVKSAISTLEDVSEYPLPIYRIDLPLETKQ
ELVKVALDPTNKNKEVOLRGIEOIAQFATSNVDLYITAKPEPKYNNIDVABGVADY
AASSYGLYQGGIIVAVADTGLDTPGRDSSMHEAPRKITLIALGRTNANDTNGHT
HVAAGVNGVSTNKKGAPOANLVPOSIMSGGLGGLPSNLQTLFSAASARLHTN
SWGAIVNGVYITDSRNVDYVRKNMTLILFAAGNBPNGGTISABGTAKNALTVGATE
NLRPSFGYADNINHVAFSSKPTKDGRIKPDVMAPGFIFLSASSLAPDSFWANH
DKRYAMGGTSMATPIVAGNVANQLRHFPMKNGITPKPILKALILAGADIGLYPN
GNQGGRTVLDKSLNAVAYNNESSLSSTGKATYSPATAGKELKISLVMSDAPASTTA
STTVNDLVDLYITAPNGTQYVGNDFTPSPNNMDEGRNVENVFINAPSGITYIEVOA
INVPGPQTFSLAIVN"

ORIGIN

Query Match 98.2%; Score 1887.8; DB 1; Length 1923;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1901; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 ATGAGAAAGAGAAAAAGGTGTTTATCTGTTTATCAGCTGCAGCGATTTTGTGACT 60
DB 1 ATGAGAAAGAGAAAAAGGTGTTTATCTGTTTATCAGCTGCAGCGATTTTGTGACT 60
QY 61 GTTGGTTAAGTAATCCATCTGCAAGTGGTGAAGAAATTTTGATCTGATTTTCAAAGA 120
DB 61 GTTGGTTAAGTAATCCATCTGCAAGTGGTGAAGAAATTTTGATCTGATTTTCAAAGA 120
QY 121 ATTGACAGCAACAATGATCTAAAGGTTTCTCCAGCAGGGCAGACTGGTCTGTCT 180
DB 121 ATTGACAGCAACAATGATCTAAAGGTTTCTCCAGCAGGGCAGACTGGTCTGTCTGCA 180
QY 181 TTTCTGTGTGAATCTGAAAAATGTGAATCTCCAAAAAGGTTTGCAGAGAAAGCTTGAAACA 240
DB 181 TTTCTGTGTGAATCTGAAAAATGTGAATCTCCAAAAAGGTTTGCAGAGAAAGCTTGAAACA 240
QY 241 GTCCCGGGCAATAATTAATCTCAATTAATCCAAATTCAGATGCAATTTTGAAGAAACA 300
DB 241 GTCCCGGGCAATAATTAATCTCAATTAATTCCAATTCAGATGCAATTTTGAAGAAACA 300
QY 301 AAACAGCAGCTGGAAGAAAAAGGGCAGAAAGTTTCTGACTACATACCTGATTTATCTTAC 360
DB 301 AAACAGCAGCTGGAAGAAAAAGGGCAGAAAGTTTCTGACTACATACCTGATTTATCTTAC 360
QY 361 ATTGTGAGATAGAGGGCGATTTAAGTCAGCAACAAGCAACATTTGAGCACTGGAATCC 420
DB 361 ATTGTGAGATAGAGGGCGATTTAAGTCAGCAACAAGCAACATTTGAGCACTGGAATCC 420
QY 421 GTGAGCCCTTAATTTCCGATATACAGAAATAGATCCCGAGCTTTTCAAAAAAGGGCAGTCA 480
DB 421 GTGAGCCCTTAATTTCCGATATACAGAAATAGATCCCGAGCTTTTCAAAAAAGGGCAGTCA 480
QY 481 GAGCTTGTAAAGCAGTGGCGCTTATACAAATGAGAAATTAAGAGAGTGCATTTAAGA 540
DB 481 GAGCTTGTAAAGCAGTGGCGCTTATACAAATGAGAAATTAAGAGAGTGCATTTAAGA 540
QY 541 GGCATGCAACAATGCAACAATTTGCAATTAAGCAATGATGTCTATATATTAACGCAAG 600
DB 541 GGCATGCAACAATGCAACAATTTGCAACAATGATGTCTATATATTAACGCAAG 600
QY 601 CCTGAGTAAAGGTGATGAATGATGCTGCGGTGAAATTTGCAAGCGGATGTGGCGAG 660
DB 601 CCTGAGTAAAGGTGATGAATGATGCTGCGGTGAAATTTGCAAGCGGATGTGGCGAG 660

QY 661 AGCAGCTACGGGTGTATGACAAGAGACAGATCGTAGCGGTGCCGATACAGGGCTGAT 720
DB 661 AGCAGCTACGGGTGTATGAGACAGAGACAGATCGTAGCGGTGCCGATACAGGGCTGAT 720
QY 721 ACAGGTGCAATGACAGTTCGATGATGAAGCTTCCGCGGAAAAATTAAGTGCATTAAT 780
DB 721 ACAGGTGCAATGACAGTTCGATGATGAAGCTTCCGCGGAAAAATTAAGTGCATTAAT 780
QY 781 GCATTGGGACCGGACGAATTAATGCCAATGATACGAATGATGATGATGATGATGATGAT 840
DB 781 GCATTGGGACCGGACGAATTAATGCCAATGATACGAATGATGATGATGATGATGATGAT 840
QY 841 TCCGATTAAGGAAAAAGGCTCCCATTAATAAGAAATGGCGCCAGCGCAATCTAGTCTTC 900
DB 841 TCCGATTAAGGAAAAAGGCTCCCATTAATAAGAAATGGCGCCAGCGCAATCTAGTCTTC 900
QY 901 CAATCTATCATGATAGCGGTGGGACCTTGAAGGACTACCTTGCAATCTGCAACCTTA 960
DB 901 CAATCTATCATGATAGCGGTGGGACCTTGAAGGACTACCTTGCAATCTGCAACCTTA 960
QY 961 TTCAAGCCAGCATACAGTGTGTGTGCGCAAAATTCATACAACTCTCGGGAGACAGCATG 1020
DB 961 TTCAAGCCAGCATACAGTGTGTGTGCGCAAAATTCATACAACTCTCGGGAGACAGCATG 1020
QY 1021 AATGGGCGCTTACACAAGATTCAGAAATGTGATGACTATGTGGCGCAAAATGATATG 1080
DB 1021 AATGGGCGCTTACACAAGATTCAGAAATGTGATGACTATGTGGCGCAAAATGATATG 1080
QY 1081 ACGATCTTTTGGTCCGGGAATGAAGACCGAACCGGCAACATCAGTGCACAGGC 1140
DB 1081 ACGATCTTTTGGTCCGGGAATGAAGACCGAACCGGCAACATCAGTGCACAGGC 1140
QY 1141 ACAGTAAATATGCAATTAACAGTGTGAGACTACGAAAACTCCGCCCAAGCTTTGGGCT 1200
DB 1141 ACAGTAAATATGCAATTAACAGTGTGAGACTACGAAAACTCCGCCCAAGCTTTGGGCT 1200
QY 1201 TATGCGGCAATTAATCAACATGATGAGCAAGTCTCTTCAAGTGTGAACGCAAGAAATGGA 1260
DB 1201 TATGCGGCAATTAATCAACATGATGAGCAAGTCTCTTCAAGTGTGAACGCAAGAAATGGA 1260
QY 1261 CGGATCAACCGGATGTCAATGCAACCGGAACTTCAATCATGACAGAAAGATCTTCTT 1320
DB 1261 CGGATCAACCGGATGTCAATGCAACCGGAACTTCAATCATGACAGAAAGATCTTCTT 1320
QY 1321 GCACCGGATTTCTCTCTTGTGGGCAACATGACATTAATTAATGATGATGATGATGATG 1380
DB 1321 GCACCGGATTTCTCTCTTGTGGGCAACATGACATTAATTAATGATGATGATGATGATG 1380
QY 1381 TCCATGGCTACACCGATCGTTGTGGAACGTGCAAGCTTGTGAGCAATTTGTGAAA 1440
DB 1381 TCCATGGCTACACCGATCGTTGTGGAACGTGCAAGCTTGTGAGCAATTTGTGAAA 1440
QY 1441 AACAGGCGATCACACCAAGCCCTTCTTAATTAAGCGGCACTGATTCGGGTGAGCT 1500
DB 1441 AACAGGCGATCACACCAAGCCCTTCTTAATTAAGCGGCACTGATTCGGGTGAGCT 1500
QY 1501 GACATCGGCTTGTGCTACCCGAACGTGAACAGAGATGGGAGCAAGTGAATGGAATAA 1560
DB 1501 GACATCGGCTTGTGCTACCCGAACGTGAACAGAGATGGGAGCAAGTGAATGGAATAA 1560
QY 1561 TCCCTGAAGCTTGTGCTATGTAAGAGATCCAGTTCTATTCACACAGCAAAAAAGGAGC 1620
DB 1561 TCCCTGAAGCTTGTGCTATGTAAGAGATCCAGTTCTATTCACACAGCAAAAAAGGAGC 1620
QY 1621 TACTGTTTACTGTCTATGTCGCGGCAAGCTTTGAAAAATCTCCGTGATGATGATGATGCC 1680
DB 1621 TACTGTTTACTGTCTATGTCGCGGCAAGCTTTGAAAAATCTCCGTGATGATGATGATGCC 1680
QY 1681 CTGCGGACGCAACTGCTTCCGTAAGCGTTGTCAATGATGATGATGATGATGATGATGAT 1740
DB 1681 CTGCGGACGCAACTGCTTCCGTAAGCGTTGTCAATGATGATGATGATGATGATGATGAT 1740
QY 1741 CCAATATGCAACAGATATGATGAAATGACTTTACTTACCTGCAATGATGATGATGAT 1800

|||||
Db 1741 CCAATGGCACAATATGTTGGAAATGACTTACTTGGCCATACATATTAATCGGAT 1800
Qy 1801 GGGCGCAATACGTAGAAAATGTTATTTATATGCAACCAAGCGGAGGTATACAA 1860
Db 1801 GGGCGCAATACGTAGAAAATGTTATTTATATGCGCCACAAAGCGGAGGTATACAA 1860
Qy 1861 GAGTACAGGCTTATATACGTACCGGTTGACCAAGACCTTCTCGTTGGCAATTGCAAT 1920
Db 1861 GAGTACAGGCTTATATACGTACCGGTTGACCAAGACCTTCTCGTTGGCAATTGCAAC 1920
Qy 1921 TAA 1923
Db 1921 TAA 1923
RESULT 9
AB046403 1920 bp DNA linear BCT 10-MAY-2002
LOCUS Bacillus sp. 9860 PROA gene for protease, complete cds.
DEFINITION AB046403
ACCESSION AB046403.2 GI:20521152
VERSION
KEYWORDS
SOURCE Bacillus sp. 9860
ORGANISM Bacillus sp. 9860
REFERENCE 1
AUTHORS Saeki,K., Okuda,M., Hatada,Y., Kobayashi,T., Ito,S., Takami,H. and
TITLE Novel oxidatively stable subtilisin-like serine proteases from
JOURNAL aliphilic Bacillus spp.: enzymatic properties, sequences, and
REFERENCES Biochem. Biophys. Res. Commun. 279 (2), 313-319 (2000)
PUBMED 1118284
2 (bases 1 to 1920)
AUTHORS Saeki,K.
TITLE Direct Submision
JOURNAL Submitted (20-JUL-2000) Katsuhisa Saeki, Kao corporation,
Biological Science Laboratory, Ichikaimachi Akabane 2606, Haga,
Tochigi 321-3497, Japan (E-mail:1387185@kcastanet.kao.co.jp,
Tel:81-285-68-7400, Fax:81-285-68-7403)
COMMENT On May 9, 2002 this sequence version replaced gi:12381938.
FEATURES
source location/Qualifiers
1..1920
/organism="Bacillus sp. 9860"
/mol_type="genomic DNA"
/strain="9860"
/db_xref="taxon:133778"
1..1920
/gene="PROA"
1..1920
/gene="PROA"
/codon_start=1
/transl_table=11
/product="protease"
/protein_id="BAB21266.2"
/db_xref="GI:20521153"
/translation="MRKKVFLSVLSAAALSTVALNPASGADRTFDLFGKIQT
DVSFSGKQRTGAALFLVSENVKLGKLETPANKLHIVQNPILBERTOK
LETTGAKLIDYIPDYAVIYIEGDDVOSKRSIEHVSVEPVLPKYRIDPOLTKGAS
LYKALALDKNNKQVQLRGIRIAQYVANSNVHYITTAPEYKVMNDVARGIYKADVA
QSSYGIYGGQIVAAVADTGLDGRNDSNHEAFRGKITALVALAGRTNNDNNGHGH
VAGSVNGATNKGMAPQANLVFQSIIDSSGIGGLPSNIQTLFSQAFSAGARIHTNS
WCAAVAGATTSRNVDYVRKNDMTILFAAGIERNGGTLISAPGAKAIVGATEN
LRPSFGSYADININVAQFSRGPETKQRIKPDVMAFGYILSRSLADSDSPMAHD
SKAVYAGTSMATPIVAGNVAOLRHFVKNRGITTPKPSILKALTAGADVLGYPNG
NGMGKRVTLDKSLNVAYNVSSSLSTSGAATTPPTATKPKLISLWSDAPASTTAS
VTLVNDLDTVITPKNCTRYVGNDFSAFPDNNMDGRNKNVNFINSQSTYITTEVQAY
NPVPGQNPNSLAIWN"

ORIGIN
Query Match 76.8%; Score 1477.2; DB 1; Length 1920;
Best Local Similarity 85.7%; Pred. No. 0;

Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
Qy 10 AAAAAAAGGTGTTTATCTGTTTATCACTGCAAGCCATTTTGTGACTGTGCTTA 69
Db 7 AAGAAAGAGGTGTTTATCTGTTTATCACTGCAAGCCATTTTGTGACTGTGCTTA 66
Qy 70 AGTAATCCATCTGCAGGTGTGCAAGAAATTTTGAATCTGGAATTCAGAAATTCAGCA 129
Db 67 AACATATCCCTGCGCTGTGTATGCAAGAACTTTGAATCTGGAATTTAAGAAATTCAGCA 126
Qy 130 ACAACTGATGCTTAAGGTTTCTTCAAGCAGGGCAGACTGTGCTCTGCTTTTGTG 189
Db 127 ACACCGAATGTCAAGTGTCTTCCAAACAGCAGCAACAGGTGCGCTGATTTCTGTG 186
Qy 190 GAATCTGAAAATGTGAACTCCCAAAAGGTTTGCAGAGAAAGCTTGAACAGTCCCGCA 249
Db 187 GAGTCTGAAAATGTGAACTCTTAAAGAAATGCTTAAAGAACTTGAACAGTCCCGCA 246
Qy 250 AATAATTAATCTCAATTTATCCAAATTCATATGACCAATTTTGAAGAAACAAACAGCAG 309
Db 247 AATAATTAATCTCAATTTATGCAATTCATATGACCAATTTTGAAGAAACAAACAGCAG 306
Qy 310 CTGAAAAAACAAGGGCAAAAGATTCTGACTACATCTGATTAATGCTTACATTTGTCAG 369
Db 307 CTAGAGACAACTGAGACAAAGATTCTGACTACATCTGATTAATGCTTACATTTGTCAG 366
Qy 370 TATGAGGGCGATTTAAGTCAAGCAACAGCACCATTGAGCAGTGAATCCGTGAGCCT 429
Db 367 TATGAGGGGGATTTGTCAGTCAAAAGCCGCTCATTTGAACAGTGAATCAGTGAGGCCA 426
Qy 430 TATTTCCGATATACGAATATAGTCCCGACTTTTCAAAAGGGCATTACAGCTTTGA 489
Db 427 TACTTGCCCAAAATACAAATTAATGATCCCGACTTTTCAAAAGGGCATTACAGCTTTGA 486
Qy 490 AAAAGCAGTGGCGCTGATACAAAGCAGAAATTAAGAGGTGAATTAAGGCAATCGAA 549
Db 487 AAGGCTTGGCGCTTATACGAAGCAACAAATTAAGAGGTGAATTAAGGCAATCGAG 546
Qy 550 CAATGCAACATTCGCAATTAAGCAATGATGCTATATATTAACGCAAAAGCTGATAT 609
Db 547 GAAATGCTCAGTACGTACCAAGCAATGATGCTATATATTAACGCAAAAGCTGATAT 606
Qy 610 AAGGTATGAATGATTTTCCGCTGATTTGCAAAAGCGAGTGTGCTCAGACAGCTAC 669
Db 607 AAGGTATGAATGATTTTCCGCTGATTTGCAAAAGCGAGTGTGCTCAGACAGCTAC 666
Qy 670 GGGTGTATGACAAGAGCAGATTCGAGCGGTGCGCATACAGGCTGTATCAGGTCCG 729
Db 667 GGTGTGTATGACAAGGCGAGATTTGCGAGTTGCGCATACAGGCTGTATCAGGAGAG 726
Qy 730 AATGACAGTTCATGATGATGAAGCTTCCGCGGAAAAATTAATGCAATTAATGATGAGGA 789
Db 727 AAGACAGTTCATGATGATGAAGCTTCCGCGGAAAAATTAATGCAATTAATGATGAGGA 786
Qy 790 CGAGCAATTAATGCAATGATTAAGATGCTATGTCATGTCATGTCGCTGCTCGATTAT 849
Db 787 CGAGCAATTAATGCAATGATTAAGATGCTATGTCATGTCATGTCGCTGCTCGATTAT 846
Qy 850 GGAACCGGCTCACTAATAAGAAATGCGGCTCAGCGGAACTGTGCTTCCAAATCTATC 909
Db 847 GGAATATGCGGCAACGAATTAAGAAATGCGGCTCAGCGGAACTGTGCTTCCAAATCTATC 906
Qy 910 ATGATATGCGGCTGAGGAGCTTGGAGAGCTTCCGATCTGCAATCTTATATGAGCCAA 969
Db 907 ATGATATGAGGAGGAGGCTTGGAGAGCTTCCGATCTGCAATCTTATATGAGCCAA 966
Qy 970 GCATACAGTGTGCTGCTGCAATTTATACAACTCTCTGGGAGCAGCAGTGAATGGGCT 1029
Db 967 GCATACAGTGTGCTGCTGCAATTTATACAACTCTCTGGGAGCAGCAGTGAATGGGCT 1026
Qy 1030 TACACAAAGATTCACAGAAATGTGATGCTATGTCGCAAAATGATATGACATCTT 1089
Db 1027 TACACAAAGATTCACAGAAATGTGATGCTATGTCGCAAAATGATATGACATCTT 1086

QY 1090 TTGCTGCGCGGGAATGAAGGACCGAA CGGCGGAACCATGATGACACAGGCA CAGCTAAA 1149
DB 1087 TTGCGCGCTGGGAATGAAGGCGGAA CGGCGGATACATGACCTGATCGGCTAAA 1146
QY 1150 AATGCAATTAACAGTCCGAGCTACGGAACCTCCGCGCAAGCTTTGGTCTTATGCGGAC 1209
DB 1147 AAGCCATTAACAGTCCGCGCAACGABAACTTGCGTCAAGCTTGGTCTCATGCAAGT 1206
QY 1210 AATATCAACCATGTGGCAAGTTCTCTTCACTGGAACCGAACAAGATGACGAGTCAAA 1269
DB 1207 AATATTAACCAAGTTCACAGTTCTCTCCGCGCGCAACAAGATGGGGAATCAAG 1266
QY 1270 CCGAGTGTATGGACCGGGAAAGTTTATATGACGAAGATCTTCTTGACACCGGAT 1329
DB 1267 CCGATGTATGCGCGCGGACGATATTTATACGAAGATCTTCTTGACACCGGAT 1326
QY 1330 TCCCTCTTGGGCGAACCATGACAGTAATATGACATATGAGTGAAGCTTCATGCT 1389
DB 1327 TCCCTCTTGGGCGAATCATGACGCAATATGCTATATGGGTGAACCTTCATGGA 1386
QY 1390 AACAGCATGTGTGCTGGAACCTGGCAAGCTTCCGAGCATTTTGTGAAAAACAGAGC 1449
DB 1387 AACAGCATGTGTGCGGGGAATGTGCAAGCTCCGAGCATTTGTGAAAAATGAGGA 1446
QY 1450 ATCAACCAAGAGCTTCTATTTAAAAAGGCGACTGATTGGCGGCGAGCTGACATCGGC 1509
DB 1447 ATCACTCTTAAAGCTTCTCTATTTGAAGCAAGCTTTGATGAGGTGCTGATGTTGA 1506
QY 1510 CTGGCTACCGGAAACGTAACCAAGATGGGGAACGAGTACATTTGATTAATCCCTGAAC 1569
DB 1507 TTGGGTTATCCGAAACGGAACCAAGATGGGCGGAGTGAACCTCGGATTAATCGTTGAAC 1566
QY 1570 GTTGCCTATGTAACGAGTCCAGTCTCTATTCACAGGCAAAAAGCGATCTGCTT 1629
DB 1567 GTTGCCTATGTAACGAGTCCAGTCCCTATCACTGACCAAAAAGCGCATATACCTTT 1626
QY 1630 ACTGCTACGCGGCAAGCTTTGAAAATCTCCCTGATATGATGGTCCGCGAGC 1689
DB 1627 ACTGCAACGCGCGGCAAGCTTTGAAAATCTCCCTGATATGATGGTCCGCGAGC 1686
QY 1690 AACATGCTCTGTAACGCTTGTCAATGATCTGGAACCTTGTATTAACGCTTCAATGAC 1749
DB 1687 ACTATGCTCTGTAACGCTTGTCAATGATTTGATTTGTGATTAACGACCAACGA 1746
QY 1750 AACAGTATGTAAGAAATGATTTATGATGACCAACGCGGAGCTTATCAATGAGGTAC 1809
DB 1747 AACAGTATGTAAGAAATGATTTATGATGACCAACGCGGAGCTTATCAATGAGGTAC 1806
QY 1810 AACAGTATGTAAGAAATGATTTATGATGACCAACGCGGAGCTTATCAATGAGGTAC 1869
DB 1807 AACAGTATGTAAGAAATGATTTATGATGACCAACGCGGAGCTTATCAATGAGGTAC 1866
QY 1870 GCTTATTAACGTAACGCTTGAACCAAGCTTCTGTTGGCAATTTGTAATTA 1923
DB 1867 GCTTATTAACGTAACGCTTGAACCAAGCTTCTGTTGGCAATTTGTAATTA 1920

RESULT 10
AR368116
LOCUS AR368116 1920 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 3 from patent US 6376227.
ACCESSION AR368116
VERSION AR368116.1 GI:34601777
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1920)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,
TITLE Shikata,S. and Nomura,M.
JOURNAL Alkaline protease
Patent: US 6376227-A 3 23-APR-2002;

Kao Corporation; Tokyo;
JPK;
FEATURES
source Location/Qualifiers
1..1920
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 76.8%; Score 1477.2; DB 6; Length 1920;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 10 AAGAAAAGGTGTTTATCTGTTTATCACTGACAGGATTTGTGACCTGTCGTTA 69
DB 7 AAGAAAAGGTGTTTATCTGTTTATCACTGACAGGATTTGTGACCTGTCGTTA 66
QY 70 AGTATCCATCTGACAGTGTGCAAGAAATTTTATCTGATTTCAAGAAATTCAGACA 129
DB 67 AACATCCCTCGGCTGATGATGCAAGACCTTTGATCTGATTTTAAAGAAATTCAGACA 126
QY 130 AACATGATCTAAAGTTTCTCCAAAGCGGCGAGACTGAGTGTGCTGCTTTTCTGAGT 189
DB 127 AACATGATCTAAAGTTTCTCCAAAGCGGCGAGACTGAGTGTGCTGCTTTTCTGAGT 186
QY 190 GAATCTGAATATGTAATCTCCAAAGGTTTGAAGAAAGCTTGAACAGTCCCGCA 249
DB 187 GAGTCTGAATATGTAATCTCTTAAAGGATTTGTAAGAAAGCTTGAACAGTCCCGCA 246
QY 250 AATATTAATCTCAATTTATCAATTTCAATGACCAATTTTGAAGAAACAAACAGCAG 309
DB 247 AATATTAATCTCAATTTATGTCATTTCAATGACCCCAATTTTGAAGAAACAAACAGCAG 306
QY 310 CTGAAAAAACAGGCGCAAGATTTCTGACTACATCTGATTTTACTTATCTTATCTGAGT 369
DB 307 CTGAGACAACTGAGAGCAAGATTTCTGACTACATCTGATTTATGATATTTGTCAG 366
QY 370 TATGAGGCGATGTTTAAATGTCAGCAACAGCACTTGAACGTTGAAATCCGTGAGCCT 429
DB 367 TATGAGGCGATGTTTAAATGTCAGCAACAGCACTTGAACGTTGAAATCCGTGAGCCT 426
QY 430 TATTTGCGATATACAGAAATGATCTCCAGCTTTTCAAAAAGGCGATGAGCTTGA 489
DB 427 TATTTGCGAAATACAGAAATGATCTCCAGCTTTTCAAAAAGGCGATGAGCTTGA 486
QY 490 AAACAGTGTGCGCTTGAATCAACAGCAAAAATTAAGAGTGCATTTAAGAGCATGAA 549
DB 487 AAACAGTGTGCGCTTGAATCAACAGCAAAAATTAAGAGTGCATTTAAGAGCATGAA 546
QY 550 CAATTCGACAAATTCGCAATTAAGATGATGTTATATTAACGCAAGCCGATAT 609
DB 547 CAATTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606
QY 610 AAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
DB 607 AAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
QY 670 GCGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
DB 667 GCGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
QY 730 AATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
DB 727 AATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
QY 790 CGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
DB 787 CGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
QY 850 GGAACGCTCTCACTAATTAAGAAATGCGCTTCAAGCAATTAATGATCTTCAATCTATC 909
DB 847 GGAACGCTCTCACTAATTAAGAAATGCGCTTCAAGCAATTAATGATCTTCAATCTATC 906
QY 910 ATGATATGCGGTGGGGAATTTGAGGACTTCTGATGATGCAAACTTATTCAGCCAA 969


```

Oy 612 GGTGATGAATGATGTCGGCGTGAATGTCAAGGCGAATGTGGCTCAAGACGACTACGG 671
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1463 AGTTTGAATGACGTGGCCCGTGGCTTGTGAAGAGACGTGGCACAAATATACCTTGG 1522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 672 GTTGTATGACGAAGACAGATCGTACCGGTTGCCATACAGGCGCTTGAATACAGTGGCA 731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1523 CTATATGACGAAGACAGATGTGACAGTTGCTGATACCTGGCTTGAATACAGAGAA 1582
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 732 TGAACGTTGATGATGAAAGCTTCGGCGGGAATTAATCTGCTATATATGCAATGGAGC 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1583 TGAACGTTGATGATGAAAGCTTCGGCGGTAAGATTAATCTGCTATATGCACTGGAGC 1642
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 792 GACGAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1643 AACGAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1702
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 852 AACCGCTCACTAATTAAGAAATGCGCTCAGCGGAATCTATCTTCCATCTATCAT 911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1703 AATATG---CTACAAATTAAGAGATGCGACCGCAAGCAATCTATCTTCAATCTATTA 1759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 912 GGATAGCGGTGGGGGACTTGGAGAGACTCTTGAATCTGCAAACTTATTCAGCCAGC 971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1760 GGATAGGTGTGAGGGCTGGAGAGACTCTGCTAATCTAACAACATTAATTCAGTCAAGC 1819
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 972 ATACAGTGTGTCGTCGACGAATTAACAATCTCTGGGAGACGAGTGAATGGGCTTA 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1820 ATATAGTGTGTCGTCGACGAATTAACAATCTATGAGGGGCTCCAGTAAACGTCCTTA 1879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1032 CAACAAGATTCAGAAATGTGATGATCTATGTCCCAAAATGATATGCAATCTCTTT 1091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1880 TACGACAGACTCTCGAAATGTTGATGATATGTAAGAAAATGATGATGATGATGATG 1939
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1092 CGTGTGCGGGAATGAAGACCGAACCGGACCATCATGTCACGACGACGACTAA 1151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1940 TCGCGCGGGAATGAAGACCGAACCGGTAACATCAGTGCACGAGAACGCAAA 1999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1152 TCGAATTAACGTCGAGCTACGGAATAACCTCCGCAAGCTTGGGCTTAATGCGGACA 1211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2000 TCGAATTAACGTCGAGCTACGGAATAACCTCCGCAAGCTTGGGCTTAATGCGGACA 2059
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1212 TATCAACATGTGCGACAGTCTCTTCACTGTGACCGCAAGAGATGACGATCAAC 1271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2060 TATTAACATGTGTGCTCAATCTCTTCAAGAGGCTCTAATGAGATGAGATTAAGGC 2119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1272 GATGTGATGCGACCGGGAACCTTCAATCATGACGAATCTTCTTTCGACCGGATTC 1331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2130 GAGCGTCATGCGACCGGTAACGTAATCTCTCTGCTAGATCATTAAGCTCAATTC 2179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1332 CTCCTTCTGGGGAACATGACGATTAATGATCATGAGGTGGAACGTCCATGCTAC 1391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2180 CTCATTTCTGGGAACATGATGATTAATGATGATGATGATGATGATGATGATGATG 2239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1392 ACCGATCGTGTGCGAAACGTGGCACAGCTTCGTGACATTTTGTGAATAACAGAGCAT 1451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2240 TCCATATGACGACGATATGTTGACAAATTAAGAGGCACTTTTGTGAATAATGAGGGGT 2299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1452 CACACCAAAAGCTTCTCTAATTAAGCGGCACTGATTCGCGTGCAGTCAATCGGCT 1511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2300 AACTCTTAAGCTTCCCTTTTAAAGCTCTTAAATTAAGAGGCTGCGGATGTTGAGCT 2359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1512 TGGGTAACCGGAACGTAACCAAGGATGAGGACGAGTGAATGGAATTAATCCTTGAACGT 1571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2360 TGGCTTTCAAATGATGACCAAGATGAGGAGAGATGAGTGAATTAATCCTTAATAT 2419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1572 TGCCTATGTAAGCAAGTCCAGTCTCTATTCACACGACCAAAAGCAGCTATCTGTTAC 1631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2420 CGCATTTGTGAATGAAAGAGCCCTTATCAACAATCAAAAAGCAACATATTCGTTAC 2479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1632 TGCCTATGTCGCGCAAGCTTGTGAATAATCTCTGATATGATGATGATGATGATGATG 1691
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2480 GGCCTCAAGCTGTGTAACCTTTAAATATCACTTGTGATGATGATGATGATGATGATG 2539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1692 AACTGCTTCGTAAGCTTGTGAATGATGATGATGATGATGATGATGATGATGATGATG 1751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 2540 GACGCGATACATCTTATGTAATGATTTAAGCTTAATGATGATGATGATGATGATG 2599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1752 ACAGATGATGAGAAATGACTTACTTCCGCTATGATGATGATGATGATGATGATGATG 1811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2600 TAAATACGTGGAATGATCTTATGACGACCGATATGATTAACAATGGGATGCGAAGAAC 2659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1812 CGTGAATAATGATTTATTAATGATGACCAAGACCGGAGATGATTAACAATGAGTACAGGC 1871
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2660 CGTGAATAATGATTTATTAATGATGACCAAGACCGGAGATGATTAACAATGAGTACAGGC 2719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1872 TTATAACGTACCGGTGACCAAGACGCTTCTGTTGGCAATTTGAATTA 1923
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2720 TTAACATGATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2771
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
E03808
LOCUS      E03808          2218 bp    DNA          linear    PAT 29-SEP-1997
DEFINITION DNA encoding alkaline protease Ya.
ACCESSION E03808
VERSION    E03808.1  GI:2172022
KEYWORDS   JP 1992197182-A/1.
SOURCE     Bacillus sp.
ORGANISM   Bacillus sp.
REFERENCE  1 (bases 1 to 2218)
AUTHORS    Toke,S., Odera,M. and Aaei,Y.
TITLE      DNA CODING ALKALINE PROTEASE YA ENZYME AND PRODUCTION OF ALKALINE
           PROTEASE YA USING THE DNA
JOURNAL     Patent: JP 1992197182-A 1 16-JUL-1992;
           LION CORP
COMMENT     OS Bacillus sp.
           PN JP 1992197182-A/1
           PD 16-JUL-1992
           PE 28-NOV-1990 JP 1990327110
           PI TOBE SEIICHI, ODERA MOTOMASU, ASAI YOSHIO
           PC C12N15/57, C11D3/386, C12N9/54, (C12N15/57, C12R1/07), (C12N9/54,
           PC C12R1/07);
           CC strandedness: Double;
           CC topology: Linear;
           CC hypothetical: No;
           CC anti-sense: No;
           CC *source: strain=Y strain;
           CC *source: clone=PBBA;
           FH Key
           FH Location/Qualifiers
           FT CDS
           FT FT 218..2125
           FT sig_peptide 218..823
           FT mat_peptide 824..2122
           FT Location/Qualifiers
           FT 1..2218
           /organism="Bacillus sp."
           /mol_type="genomic DNA"
           /db_xref="taxon:1409"

ORIGIN
Query Match 48.7%; Score 936.2; DB 6; Length 2218;
Best Local Similarity 69.4%; Pred. No. 2.1e-228;
Matches 1335; Conservative 0; Mismatches 573; Indels 15; Gaps 4;

Oy 1 ATGAGAAAGAGAAAGAGTGTATCTGTTTATCTGATGATGATGATGATGATGATGATG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218 ATGAGAGGAGAAAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 61 GTTGGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 GTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 121 ATTGAGCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 332 GTGAAAAGTTTAGAAAATGCTA---CTTTGTTAAACCGATAAGTAGCGGTGAGCATCC 388
 Qy 181 TTTTCGTGTTGAAATCTGAAAATGTGAAATCTCCAAAAGTTTGAGAAAGCTTGAAACA 240
 Db 389 TTTCTAGTAGAATACGGAATAATTAATATCTTAAAGGTATTCAAAAGAGCTAGAACGA 448
 Qy 241 GTCCGGCAATATATTAATCTCAATATTAATCAATTCATGAGCAATTTTAGAAGAAACA 300
 Db 449 GTACAGAGATTAACGAACTCTACATCTGACATTTTACTGACCAATTTTCAGAGAGAG 508
 Qy 301 AAACAGAGCTGGAAGAAACAGGGGCAAGATTCCTGACTACATACCTGATTATGCTTAC 360
 Db 509 CCAAAAGAGTTAGAGTCTAGAGATGATGATTTAGATTAGTTGCTCAATATATGTTTT 568
 Qy 361 ATTTCGAGTATGAGGGCGATGTTAAGTACGACACAGCAATTTGAGCACTGGAATCC 420
 Db 569 ATTTGTCAGTAT---AGTGTGCTACAAAATAATTAAGTACTTTACTCTTTGAGAAC 625
 Qy 421 GTGAGCTTATTTGGCCGATATACAGATPAGATCCCGAGCTTTTCAGAAAAGGGCATCA 480
 Db 626 GTACAACTTTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 685
 Qy 481 GAGCTTGTAAAGCAGTGGCGCTGTATACAAAGCAAGAAATTAAGAGGTGCAATTAAGA 540
 Db 686 CAGCTTGTTCAGCGGTATTTTAAATACAAAACAGAAATTAATAATTAATTAATTAATTA 745
 Qy 541 GGCAATCGAACAATTCGCACAAATTCGCAATTAAGCAATGATGCTATATTTACGGCAAG 600
 Db 746 GATTAGATGAGATGTTCAATATCTGCAATTAATGATGCTTTATATATCAACCAAG 805
 Qy 601 CTTGAGTAAAGGTATGATGATGTTGGCGCTGGAATTTGCAAGCGGATGCTGACG 660
 Db 806 CCCGAGTATGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 865
 Qy 661 AGCAGCTACCGGTTGATGAGCAAGCAAGATCTGAGCGGTTGCCGATACAGGCTTAT 720
 Db 866 AACATTTACGATATATGAGCAAGCTCACTAGTTGAGTACGAGCAAGGCTTATGAT 925
 Qy 721 ACAGTCCCAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 Db 926 ACAGTCCCAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 985
 Qy 781 GCATTGGAGCGGACCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
 Db 986 GCGTTAGGAAGAACTAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGAT 1045
 Qy 841 TCCGTTATGAGAAACGCTCCACTAATTAAGGAATGCGGCTCAGGCGAATCTTACTTTC 900
 Db 1046 TCTGTTCTTTGTAATG---CTTTAAATTAAGGAATGCGCTCCGCAAGCTTAATCTTTC 1102
 Qy 901 CAATCTATCATGATGAGGAGGTGGGGGACTTGAGGAGCACTTTCGAAATCTGCAACCTTA 960
 Db 1103 CAATCTATTAATGATGAGGAGGAGGAGGATTTAGGTGCTTAACATTCGAACTTAATTA 1162
 Qy 961 TTCAGCCAAAGATACAGTGTCTGTGTCAGAAATTCATACAACTCTCTGGGAGCAGCAGT 1020
 Db 1163 TTTAGTCAAGCTTGAATGCTGAGCAAGAAATTAATTAATCTTTGGGAGCCCCAGTA 1222
 Qy 1021 AATGGGCTTACACAAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1080
 Db 1223 AATGAGGCTACATGCTTACCTGAGCAAGTGAATGATGATGATGATGATGATGATGATGAT 1282
 Qy 1081 AGGATCTTTTTCGTCGCGGGGATGAAGACCGAACCGGCGAACATGATGAGCAACAGGC 1140
 Db 1283 AGGATCTTTTTCGTCGCGGGGATGAAGACCGAACCGGCGAACATGATGAGCAACAGGC 1342
 Qy 1141 ACAGCTAAATTAATGCAATGATGAGTACGAGAACTCCGCGCAAGCTTTGAGTCT 1200
 Db 1343 ACAGGAAATTAATGCAATGATGAGTACGAGAACTCCGCGCAAGCTTTGAGTCT 1402
 Qy 1201 TATGCGGCAATTAATGATGAGTACGAGTCTTCTTCACTGAGCCGACAAAGGATGA 1260
 Db 1403 ATAGCAGATTAACCAATCATATGCAATTTTCATGAGAGGAGCTACGAGGAGATGA 1462

Qy 1261 CGGATCAAAACGGATGTCAATGCAACCGGAAAGCTTCAATACATACGAAATCTTCTCT 1320
 Db 1463 CGAATTAAGCTTGAAGTAAAGCTTCTGAAACATTTATTTATCAGACAGTCTTCTCTTA 1522
 Qy 1321 GCACCGGATCTCTCTCTGAGCGGAAACGATACAGTAAATTAATGCAATACAGTGGTGAAC 1380
 Db 1523 GCTTCAGCTCTTCTCTTGGGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1582
 Qy 1381 TCCATGCTTACACCGATGCTTGTGAAACGTTGCAACGCTTCTGAGCAATTTGTGAAA 1440
 Db 1583 TCCATGCTTACACCGATGCTTGTGAAACGTTGCAACGCTTCTGAGCAATTTGTGAAA 1642
 Qy 1441 AACAGAGCTACACCAAGCTTCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTA 1500
 Db 1643 AATAGAGTATTAATCTTAAAGCTTCTTAAATTAATTAATTAATTAATTAATTAATTA 1702
 Qy 1501 GACATCGGCTTGTGCTACCGAAGGATACCAAGATGAGGAGGAGGAGGAGGAGGAGGAG 1560
 Db 1703 GATGTTGTTTAAAGATATCTTAAAGGATACCAAGGATGAGGAGGAGGAGGAGGAGGAG 1762
 Qy 1561 TCCCTGAACGTTGCTATGAGCAAGTCCAGTCTCTATCCACAGCCAAAGAGGAG 1620
 Db 1763 TCGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1822
 Qy 1621 TACTGTTTACTGCTACTGCGGCAAGCTTTGAAATCTCCCTGATATGATGATGATGAT 1680
 Db 1823 TATTCTTTCGAAGCAAGCGGATTAACCTTTAAATTAATTAATTAATTAATTAATTAAT 1882
 Qy 1681 CCTGAGAGCAACATGCTTCCGTAACGCTTGTCAATGATCTGACCTTGTCAATACGCT 1740
 Db 1883 CCTGAGAGCAACATGCTTCCGTAACGCTTGTCAATGATCTGACCTTGTCAATACGCT 1942
 Qy 1741 CCAATGAGCAACATGATGAGAAATGATCTTAACTTCCGCAATCAATGATTAATGAT 1800
 Db 1943 CCGAATGAGCAACATGATGAGAAATGATCTTAACTTCCGCAATCAATGATTAATGAT 2002
 Qy 1801 GCGCGCAATTAACGTAAGAAATGATTTTAAATTAATTAATTAATTAATTAATTAATTA 1860
 Db 2003 GGTGCAACCAATGTTGAGAAAGTATTTAAATTAATTAATTAATTAATTAATTAATTA 2062
 Qy 1861 GAGGTACAGGCTTAAAGTACGTTCCGTTGAGCAACAGACCTTCTGTTGGCAATGTTGAT 1920
 Db 2063 GAGGTACAGGCTTAAAGTACGTTCCGTTGAGCAACAGACCTTCTGTTGGCAATGTTGAT 2122
 Qy 1921 TAA 1923
 Db 2123 TAA 2125

Search completed: April 8, 2006, 06:48:17
 Job time : 9427.85 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 00:55:07 ; Search time 2984.33 Seconds
(without alignments)
4294.503 Million cell updates/sec

Title: US-10-784-870-5

Perfect score: 1923

Sequence: 1 atgagagaagaagaagaagt.....cgttgcaattgtgaataa 1923

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries.

Database :

1: N_Geneseq_21: *
2: geneseqn1980s: *
3: geneseqn1990s: *
4: geneseqn2000s: *
5: geneseqn2001as: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *
14: geneseqn2005s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1923	100.0	1923	2	AAx37278 Bacillus
2	1923	100.0	1923	12	Adm40772 DNA encod
3	1923	100.0	1923	13	Adm14428 Bacillus
4	1923	100.0	1923	13	Adt49605 Bacillus
5	1919.8	99.8	1923	2	AAx37279 Bacillus
6	1477.2	76.8	1920	2	AAx37277 Bacillus
7	1305	67.9	1305	11	ADY33779 Bacillus
8	1305	67.9	1305	12	ADL25803 Bacillus
9	1166.4	60.7	3003	2	AAV82382 Bacillus
10	737.2	38.3	1299	2	AAQ27516 Alkali-pr
11	53.6	2.8	2000	8	ADA71938 Rice gene
12	50	2.6	2000	8	ADA71938 Rice gene
13	48	2.5	54	11	ADY33797 Bacillus
14	48	2.5	54	12	ADL25809 Bacillus
15	45	2.3	415	8	ABO82958 Human lun
16	41.8	2.2	464	13	ACP89679 Human SIR
17	41.8	2.2	794	4	AAI96559 Human neu
18	41.8	2.2	1095	3	AAC59782 Human sec
19	41.8	2.2	1095	4	AAH33984 Human col

C 20	41.8	2.2	1297	6	ABK35920	Abk35920 cDNA sequ
C 21	41.8	2.2	5053	4	AAI58702	Aai58702 Human pol
C 22	41.8	2.2	5053	5	ADQ98920	Adq98920 DNA encod
C 23	41.8	2.2	5053	9	ADB48680	Adb48680 Novel hum
C 24	41.8	2.2	5396	4	AAK52078	Aak52078 Human pol
C 25	41.8	2.2	8195	4	AAK81200	Aak81200 Human imm
C 26	41.8	2.2	8195	4	AAI06335	Aai06335 Human rep
C 27	41.8	2.2	9035	6	ABK89430	Abk89430 Human tum
C 28	41.8	2.2	9229	8	ABX34821	Abx34821 Human tmd
C 29	41.8	2.2	27377	6	ABK89433	Abk89433 Human tum
C 30	40.4	2.1	2000	11	ACL37108	ACL37108 Rice stre
C 31	39.8	2.1	6843	6	ABK82368	Abk82368 DNA trans
C 32	39.4	2.0	4414	4	ABL06106	AbL06106 Drosophi1
C 33	39.2	2.0	4590	5	AAH24065	Aah24065 Yeast AOD
C 34	39	2.0	7380	13	ADT66571	Adt66571 Murine Lu
C 35	39	2.0	7380	14	ADX26026	Adx26026 Mouse cel
C 36	39	2.0	110000	6	ABA03041	ABA03041 cont. (16 o
C 37	39	2.0	342748	14	ADZ13793	Adz13793 Human can
C 38	38.8	2.0	766	6	ABQ45295	Abq45295 Oligonuc1
C 39	38.8	2.0	2000	11	ACL35887	ACL35887 Rice stre
C 40	38.8	2.0	1483	14	ABR70208	ABr70208 Helicobac
C 41	38	2.0	67253	14	ABR61178	ABr61178 Human GPR
C 42	37.8	2.0	2835	2	AAI08141	Aai08141 Hyperther
C 43	37.6	2.0	4765	2	AAI08132	Aai08132 Protease
C 44	37.6	2.0	4765	2	AAI85670	Aai85670 Pyrococu
C 45	37.6	2.0	4765	2	AAI85670	Pyrococu

ALIGNMENTS

RESULT 1	AAx37278	standard; DNA; 1923 BP.
ID	AAx37278	
AC	AAx37278;	
DT	20-MAR-2003 (revised)	
DT	21-JUL-1999 (first entry)	
XX		
DE	Bacillus alkaline protease encoding DNA.	
KW	alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;	
KW	washing composition; oxidizing agent; ss.	
XX		
OS	Bacillus sp.	
PN	MO9918218-A1.	
XX		
PD	15-APR-1999.	
XX		
PF	07-OCT-1998; 98WC-JP004528.	
XX		
PR	07-OCT-1997; 97JP-00274570.	
XX		
PA	(KAOS) KAO CORP.	
XX		
PI	Takaiwa M, Okuda M, Saeki K, Kubota H, Hitoi J, Kageyama Y;	
PI	Shiketa S, Nomura M;	
DR	WPI, 1999-287736/27.	
DR	P-PSDB; AAY17088, AAY17090.	
XX		
PT	Alkali protease from Bacillus used in washing powders.	
XX		
PS	Disclosure; Page 58-63; 71pp; Japanese.	
CC	The invention relates to alkaline proteases produced by strains of	
CC	Bacillus. The proteases ability to digest casein is not inhibited by	
CC	oleic acid and they have a high stability to oxidizing agents. The	
CC	alkaline protease of the invention has the following properties: (a) it	
CC	is active over the pH range 4-13 and has at least 80% of its optimum	
CC	activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is	


```

Db 1141 ACAGCTAAATGCAATTAACAGTCGAGCTACGAAACCTCCGCCAAGCTTTGGTCT 1200
QY 1201 TATGCGGACAAATATCAACATGAGCAAGTTCTTCTTCAAGTGGACCGCAAGATATGA 1260
Db 1201 TATGCGGACAAATATCAACATGAGCAAGTTCTTCTTCAAGTGGACCGCAAGATATGA 1260
QY 1261 CGGATCAAAACCGGATGTCATGAGCAACCGGAACTTCAATATCAGCAAGATCTTCTT 1320
Db 1261 CGGATCAAAACCGGATGTCATGAGCAACCGGAACTTCAATATCAGCAAGATCTTCTT 1320
QY 1321 GCAACCGGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1380
Db 1321 GCAACCGGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1380
QY 1381 TCCATGCGTACACCGGATGTCATGAGCAACCGGAACTTCAATATCAGCAAGATCTTCTT 1440
Db 1381 TCCATGCGTACACCGGATGTCATGAGCAACCGGAACTTCAATATCAGCAAGATCTTCTT 1440
QY 1441 AACAGAGGATCAACACCAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1500
Db 1441 AACAGAGGATCAACACCAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1500
QY 1501 GACATGCGCTTGGCTACCGGAACTTCAATATCAGCAAGATCTTCTTCTTCTTCTTCTT 1560
Db 1501 GACATGCGCTTGGCTACCGGAACTTCAATATCAGCAAGATCTTCTTCTTCTTCTTCTT 1560
QY 1561 TCCCTGAAAGCTTGGCTACCGGAACTTCAATATCAGCAAGATCTTCTTCTTCTTCTTCTT 1620
Db 1561 TCCCTGAAAGCTTGGCTACCGGAACTTCAATATCAGCAAGATCTTCTTCTTCTTCTTCTT 1620
QY 1621 TACTGCTTACTGCTACCTGCGGCAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1680
Db 1621 TACTGCTTACTGCTACCTGCGGCAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1680
QY 1681 CTTGCGAGACAACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1740
Db 1681 CTTGCGAGACAACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1740
QY 1741 CCAATGCGACAACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1800
Db 1741 CCAATGCGACAACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1800
QY 1801 GCGCGCAATTAACGTAAGAAATGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
Db 1801 GCGCGCAATTAACGTAAGAAATGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
QY 1861 GAGGTACAGGCTTATTAACGTAAGAAATGTAATTAATTAATTAATTAATTAATTAATTAAT 1920
Db 1861 GAGGTACAGGCTTATTAACGTAAGAAATGTAATTAATTAATTAATTAATTAATTAATTAAT 1920
QY 1921 TAA 1923
Db 1921 TAA 1923

```

```

RESULT 3
ADSL14428
ID ADS14428 standard; DNA; 1923 BP.
XX AC
XX ADS14428;
XX
XX 30-DEC-2004 (first entry)
XX
XX Bacillus alkaline protease KP43 gene SEQ ID NO.2.
XX
XX protease; enzyme; alkaline protease; laundry detergent; KP43; db; gene.
XX
XX Bacillus sp.; KSM-KP43.
XX
XX Key Location/Qualifiers
XX 1..1923
XX CDS /*tag= b
XX FT

```

```

FT sig_peptide /product= "alkaline protease"
FT 1..618
FT mat_peptide /*tag= a
FT 619..1920
FT /*tag= c
PN EP146962-A1.
PD 13-OCT-2004.
PF 08-APR-2004; 2004EP-00008604.
PR 10-APR-2003; 2003JP-00106709.
PA (KAOS ) KAO CORP.
PI Okuda M, Sato T, Takimura Y, Sumitomo N, Kobayashi T;
DR WPI; 2004-711313/70.
DR P-PSDB; ADS14437, ADS14427.
PT New engineered alkaline protease, useful particularly in laundry
FT detergents, comprising specified amino acids at particular positions.
PS Disclosure; SEQ ID NO 2; 31pp; English.
CC The invention relates to a novel alkaline protease. The new alkaline
CC protease comprises an amino acid sequence in which one or more amino acid
CC residues selected from those located at 7 specific positions within
CC 15 (histidine), position 16 (threonine or glutamine), position 166
CC (glycine), position 167 (valine), position 187 (serine), position 346
CC (arginine), and position 405 (aspartic acid). The alkaline protease is
CC useful in industry particularly in laundry detergents, but also e.g. in
CC fibre modifying agents, leather processing agents, cosmetic compositions,
CC bath additives, food-modifying agents, and pharmaceuticals. The present
CC sequence encodes the wild-type Bacillus sp. KSM-KP43 alkaline protease,
CC KP43.
XX
XX SEQ Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 1923; DB 13; Length 1923;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAAGAAAGAAAGAGTGTGTTTATCTGTTTATCAGTGCAGGATTTTGGACT 60
Db 1 ATGGAAGAAAGAAAGAGTGTGTTTATCTGTTTATCAGTGCAGGATTTTGGACT 60
QY 61 GTTGCCTTAAGTATTCATCTGACAGTGTGCAAGAAATTTGATCTGATTTCAAGGA 120
Db 61 GTTGCCTTAAGTATTCATCTGACAGTGTGCAAGAAATTTGATCTGATTTCAAGGA 120
QY 121 ATTCAACAACACTGATGCTAAAGTCTTCTCAAGACAGGAGAGTGTGCTGCT 180
Db 121 ATTCAACAACACTGATGCTAAAGTCTTCTCAAGACAGGAGAGTGTGCTGCT 180
QY 181 TTTCTGTGGAATTTGAAATGTGAACTCCCAAAAGTTTGCAGAAAGCTTGAACA 240
Db 181 TTTCTGTGGAATTTGAAATGTGAACTCCCAAAAGTTTGCAGAAAGCTTGAACA 240
QY 241 GTCCGCGCAATTAATTAATCTCATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
Db 241 GTCCGCGCAATTAATTAATCTCATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
QY 301 AACAGAGCTGCAAGAAACAGGCGCAAGATTTCTGACATACCTGATTAATGCTTAC 360
Db 301 AACAGAGCTGCAAGAAACAGGCGCAAGATTTCTGACATACCTGATTAATGCTTAC 360
QY 361 ATTGTGAGTATGAGGCGGATTTAATGTCAGCAACAGCACTTGGACAGTGGATCC 420
Db 361 ATTGTGAGTATGAGGCGGATTTAATGTCAGCAACAGCACTTGGACAGTGGATCC 420

```

QY 421 GTGAGCGCTTATTTGCGATATACAGAAATAGATCCCAAGCTTTTCACAAAAAGGGGACATCA 480
 DB 421 GTGAGCGCTTATTTGCGATATACAGAAATAGATCCCAAGCTTTTCACAAAAAGGGGACATCA 480
 QY 481 GAGCTTGTAAAGCAAGTGGCGCTTGTATACAAAGCAGAAAAATTAAGAGTGCATTTAGA 540
 DB 481 GAGCTTGTAAAGCAAGTGGCGCTTGTATACAAAGCAGAAAAATTAAGAGTGCATTTAGA 540
 QY 541 GGGATGGAACAAATCCGACAAATTCGCAATTAAGCAATGATGTCTATATATTAAGCGCAAG 600
 DB 541 GGGATGGAACAAATCCGACAAATTCGCAATTAAGCAATGATGTCTATATATTAAGCGCAAG 600
 QY 601 CCTGAGTATTAAGGTGATGAAATGATGTTGGCGGTGAAATTTGCAAAAGCGATGTGGCTCAG 660
 DB 601 CCTGAGTATTAAGGTGATGAAATGATGTTGGCGGTGAAATTTGCAAAAGCGATGTGGCTCAG 660
 QY 661 AGCAGCTACGGGTTGTATGACAAAGACAGATCGTACGGTTCGCCATACAGGGCTTGAAT 720
 DB 661 AGCAGCTACGGGTTGTATGACAAAGACAGATCGTACGGTTCGCCATACAGGGCTTGAAT 720
 QY 721 ACAGTTCGCAATGACAGTTGATGATGAAACCTTCGCGGGAAAAATTAATCTGCATTAAT 780
 DB 721 ACAGTTCGCAATGACAGTTGATGATGAAACCTTCGCGGGAAAAATTAATCTGCATTAAT 780
 QY 781 GCAATGAGGACGAGCAATTAATGCAATGATGCAATGATGCAATGATGCAATGATGCAATGAT 840
 DB 781 GCAATGAGGACGAGCAATTAATGCAATGATGCAATGATGCAATGATGCAATGATGCAATGAT 840
 QY 841 TCCGATTTAGAAACGGCTCCACTAAATTAAGAAATGAGCGCTCAGCGAATCTAGTCTTC 900
 DB 841 TCCGATTTAGAAACGGCTCCACTAAATTAAGAAATGAGCGCTCAGCGAATCTAGTCTTC 900
 QY 901 CAATCTATCAATGATAGCGGTGGGGAATTTGAGAGCACTACCTTCGATCTGCAACCTTA 960
 DB 901 CAATCTATCAATGATAGCGGTGGGGAATTTGAGAGCACTACCTTCGATCTGCAACCTTA 960
 QY 961 TTCAGGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 DB 961 TTCAGGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 QY 1021 AATGAGCTTACCAACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1080
 DB 1021 AATGAGCTTACCAACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1080
 QY 1081 AGGATCTTTTGGCTCCGCGGAAATGAGACGCGAGCGGAAACATCAAGTGCACCAAGC 1140
 DB 1081 AGGATCTTTTGGCTCCGCGGAAATGAGACGCGAGCGGAAACATCAAGTGCACCAAGC 1140
 QY 1141 ACAGCTTAAATGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 1141 ACAGCTTAAATGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 1201 TATGCGGACAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 DB 1201 TATGCGGACAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 QY 1261 CGGATCAAAACGGATGTCAATGACCGGGAACGTTTCAATCAAGCAAGATCTTCTCTT 1320
 DB 1261 CGGATCAAAACGGATGTCAATGACCGGGAACGTTTCAATCAAGCAAGATCTTCTCTT 1320
 QY 1321 GCAACCGGATCT 1380
 DB 1321 GCAACCGGATCT 1380
 QY 1381 TCCATGCTACACCGATCGTGTGGAACGTTGCAAGTGCAGATTTTGTGAAA 1440
 DB 1381 TCCATGCTACACCGATCGTGTGGAACGTTGCAAGTGCAGATTTTGTGAAA 1440
 QY 1441 AACAGAGGATCAACCAAGCTTCTATTAATTAAGCGGCACTGATGCGGTGACGT 1500
 DB 1441 AACAGAGGATCAACCAAGCTTCTATTAATTAAGCGGCACTGATGCGGTGACGT 1500
 QY 1501 GACATGGCGCTTGGCTACCGGAACGTTAACCAAGATGGGAGAGATGACATTTGATAAA 1560

DB 1501 GACATGGCGCTTGGCTACCGGAACGTTAACCAAGATGGGAGAGATGACATTTGATAAA 1560
 QY 1561 TCCCTGAAGCTTGCCTATGTAAGACAGTCCAGTCTCTATTCACCAAGCCAAAAGCGACG 1620
 DB 1561 TCCCTGAAGCTTGCCTATGTAAGACAGTCCAGTCTCTATTCACCAAGCCAAAAGCGACG 1620
 QY 1621 TACTGTTTACTGCTACTGCGCGGCAAGCTTTGAAAATCTCCCTGTAATGCTGATGCC 1680
 DB 1621 TACTGTTTACTGCTACTGCGCGGCAAGCTTTGAAAATCTCCCTGTAATGCTGATGCC 1680
 QY 1681 CCTGGAGGACAACTGCTCCGTAAGCTTTGCAATGATGCAATGATGCAATGATGCAATGAT 1740
 DB 1681 CCTGGAGGACAACTGCTCCGTAAGCTTTGCAATGATGCAATGATGCAATGATGCAATGAT 1740
 QY 1741 CCNAATGGAACAAGATGATGAGAAATGACTTACTTTCGCCATACATGATTAATGAGAT 1800
 DB 1741 CCNAATGGAACAAGATGATGAGAAATGACTTACTTTCGCCATACATGATTAATGAGAT 1800
 QY 1801 GGCAGCAATPACGTAGAAAATGATTAATGACCAACCAAGCGGACGTATACATTT 1860
 DB 1801 GGCAGCAATPACGTAGAAAATGATTAATGACCAACCAAGCGGACGTATACATTT 1860
 QY 1861 GAGGTACAGGCTTATACGTACCGGTTGACCAAGACCTTCCTGTTGCAATTTGCAAT 1920
 DB 1861 GAGGTACAGGCTTATACGTACCGGTTGACCAAGACCTTCCTGTTGCAATTTGCAAT 1920
 QY 1921 TAA 1923
 DB 1921 TAA 1923

RESULT 4
 ADT49605
 ID ADT49605 standard; DNA; 1923 BP.
 AC ADT49605;
 AC 30-DEC-2004 (first entry)
 DT 30-DEC-2004 (first entry)
 DB Bacillus alkaline protease encoding DNA.
 KW Alkaline protease; detergent; fiber modification; leather processing;
 KW cosmetic; bath additives; food-modification; pharmaceutical; gene; ds.
 OS Bacillus sp. KSM-KP43.
 FH Key
 FT CDS Location/Qualifiers
 FT 1..1923
 FT /tag= b
 FT /product= "alkaline protease"
 FT /note= "SEQ ID 10"
 FT sig_peptide 1..618
 FT /tag= a
 FT /note= "signal peptide"
 FT 619..1923
 FT /tag= c
 FT /product= "mature alkaline protease"
 FT /note= "SEQ ID 1"
 CDS
 EP1466970-A1.
 PD 13-OCT-2004.
 PF 08-APR-2004; 2004EP-00008605.
 PR 10-APR-2003; 2003JP-00106708.
 PA (KAOS) KAO CORP.
 PI Sato T, Okuda M, Koyama S, Izawa Y, Kobayashi T.
 XX WPI; 2004-711317/70.

DR P-PSDB; ADT49604, ADT49613.

XX New engineered alkaline protease with improved activity and thermal
PT stability, useful particularly in detergents such as laundry detergents.

XX Example 2; SEQ ID NO 2, 35pp; English.

XX The invention relates to an alkaline protease and its encoding gene. The
CC alkaline protease can be expressed by standard recombinant methodology.
CC The alkaline protease is useful in the industry particularly in
CC detergents such as laundry detergents, but also in fiber modifying
CC agents, leather processing agents, cosmetic compositions, bath additives,
CC food-modifying agents, and pharmaceuticals. The enzyme has good activity
CC and thermal stability. The present sequence represents a DNA encoding an
XX alkaline protease from *Bacillus* sp. KSM-RP43.

XX Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;

Query Match 100.0%; Score 1923; DB 13; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAAAGAGAAAAAGGTGTTTATCTGTTTATCAGCTGCAGCGATTTTGTGACT 60
DB 1 ATGAGAAAGAGAAAAAGGTGTTTATCTGTTTATCAGCTGCAGCGATTTTGTGACT 60
QY 61 GTTGCGTTAAGTAATCCATCTGCAGGTGGTGAAGAAATTTTGTGATTTTCAAGGA 120
DB 61 GTTGCGTTAAGTAATCCATCTGCAGGTGGTGAAGAAATTTTGTGATTTTCAAGGA 120
QY 121 ATTGAGACAACAATGATGCTAAAGTTTCTCCAGCAGGGGAGACTGGTCTGTCT 180
DB 121 ATTGAGACAACAATGATGCTAAAGTTTCTCCAGCAGGGGAGACTGGTCTGTCT 180
QY 181 TTTCTGTGTGAATCTGAAAAATGTGAACCTCCAAAAAGTTTGACAGAAAGCTTGAACA 240
DB 181 TTTCTGTGTGAATCTGAAAAATGTGAACCTCCAAAAAGTTTGACAGAAAGCTTGAACA 240
QY 241 GTTCCGGGCAATTAATTAATCTCAATTAATCCAAATTCAGATGACCAATTTTGAAGAAACA 300
DB 241 GTTCCGGGCAATTAATTAATCTCAATTAATCCAAATTCAGATGACCAATTTTGAAGAAACA 300
QY 301 AAACAGCAGCTGGAAGAAAAACAGGGGCAAAAGATTCTGACTACATGCTGATTTATCTTAC 360
DB 301 AAACAGCAGCTGGAAGAAAAACAGGGGCAAAAGATTCTGACTACATGCTGATTTATCTTAC 360
QY 361 ATTGTGAGATGAGAGGGCGATGTTAAGTCAGCAACAAGCACCATTGAGCAGTGAATCC 420
DB 361 ATTGTGAGATGAGAGGGCGATGTTAAGTCAGCAACAAGCACCATTGAGCAGTGAATCC 420
QY 421 GTTGAGCCTTATTTGCGGATATACAGAAATAGATCCCGAGCTTTTCAAAAAGGGGCATCA 480
DB 421 GTTGAGCCTTATTTGCGGATATACAGAAATAGATCCCGAGCTTTTCAAAAAGGGGCATCA 480
QY 481 GAGCTTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAAAATTAAGAGTGCATTAAGA 540
DB 481 GAGCTTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAAAATTAAGAGTGCATTAAGA 540
QY 541 GGCACTGAAACAATGTCACATTTCCGATTAAGCAATGATGTCTATATTAATTAACGGCAAG 600
DB 541 GGCACTGAAACAATGTCACATTTCCGATTAAGCAATGATGTCTATATTAATTAACGGCAAG 600
QY 601 CCTGAGTAAAGGTGATGATGATGTTGCGGTGGAATTTGCAAAAGCGGAGTGGCTCAG 660
DB 601 CCTGAGTAAAGGTGATGATGATGTTGCGGTGGAATTTGCAAAAGCGGAGTGGCTCAG 660
QY 661 AGCAGCTTACGGGTTGTATGACAAGAGACAGATCTGAGCGTTGCCGATACAGGGCTTGAT 720
DB 661 AGCAGCTTACGGGTTGTATGACAAGAGACAGATCTGAGCGTTGCCGATACAGGGCTTGAT 720
QY 721 ACAGGTGCCAATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACAGGTGCCAATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780

QY 781 GCATTGGAGCGGACGAATTAATCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCATTGGAGCGGACGAATTAATCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGATTAAGGAAAGGCTCCATTAATTAAGAAATGGGCGCTCAGGCGAAATCTAGTCTTC 900
DB 841 TCCGATTAAGGAAAGGCTCCATTAATTAAGAAATGGGCGCTCAGGCGAAATCTAGTCTTC 900
QY 901 CAATCTATCATGAT 960
DB 901 CAATCTATCATGAT 960
QY 961 TTGACGCAAGATACAGTGTGTGTGTCGAAATTCATACAAATCTCTGGGAGCAGCAGTG 1020
DB 961 TTGACGCAAGATACAGTGTGTGTGTCGAAATTCATACAAATCTCTGGGAGCAGCAGTG 1020
QY 1021 AATGGGCTTACACAAGATTTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATGGGCTTACACAAGATTTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ACGATCTTTTGTGCTGCGGGAATGAAGACCGAAGCGGGAACCATCATGTCACACAGGC 1140
DB 1081 ACGATCTTTTGTGCTGCGGGAATGAAGACCGAAGCGGGAACCATCATGTCACACAGGC 1140
QY 1141 ACAGCTAAATAATGCAATTAACAGTGTGAGCTACGAAAACTCCGCCAAGCTTTGGGTCT 1200
DB 1141 ACAGCTAAATAATGCAATTAACAGTGTGAGCTACGAAAACTCCGCCAAGCTTTGGGTCT 1200
QY 1201 TATGCGGCAATTAATCAACATGATGAGCAGTTCTCTTCAAGTGTGACCGAACAAGATGGA 1260
DB 1201 TATGCGGCAATTAATCAACATGATGAGCAGTTCTCTTCAAGTGTGACCGAACAAGATGGA 1260
QY 1261 CGATTCAAAACCGGATGTGATGAGCAACGGGAAGCTTCAATCATGAGAAAGATCTTCTT 1320
DB 1261 CGATTCAAAACCGGATGTGATGAGCAACGGGAAGCTTCAATCATGAGAAAGATCTTCTT 1320
QY 1321 GCACCGGATTTCTCTCTTGTGGGGCAACATGACATGATTAATGATACATGAGTGAACG 1380
DB 1321 GCACCGGATTTCTCTCTTGTGGGGCAACATGACATGATTAATGATACATGAGTGAACG 1380
QY 1381 TCCATGGCTACACCGATGTTGTGTGAAAGGTGACAGGCTGTGTGAGCAATTTGTGAAA 1440
DB 1381 TCCATGGCTACACCGATGTTGTGTGAAAGGTGACAGGCTGTGTGAGCAATTTGTGAAA 1440
QY 1441 AACAGAGCATCACACCAAGCCTTCTCTATTAAGCGGCACTGATTTGCCGCTGACCT 1500
DB 1441 AACAGAGCATCACACCAAGCCTTCTCTATTAAGCGGCACTGATTTGCCGCTGACCT 1500
QY 1501 GACATCGGCTTGTGCTACCGGAACGGTAAACCAAGAGTGGGAGCGAGTGAATGGATTA 1560
DB 1501 GACATCGGCTTGTGCTACCGGAACGGTAAACCAAGAGTGGGAGCGAGTGAATGGATTA 1560
QY 1561 TCCCTGAACGTTGCTATGTAAGAGAGTCCAGTTCTTATTCACACGACCAAAAAGGACG 1620
DB 1561 TCCCTGAACGTTGCTATGTAAGAGAGTCCAGTTCTTATTCACACGACCAAAAAGGACG 1620
QY 1621 TACTGTTTACTGCTACGTCGCGAAGCCTTTGAAAATCTCCGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
DB 1621 TACTGTTTACTGCTACGTCGCGAAGCCTTTGAAAATCTCCGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
QY 1681 CTTGCGAGCAACAATGCTTCTGTAAGCGTTGTCAATGATGAGACCTTGTCAATTAACGCT 1740
DB 1681 CTTGCGAGCAACAATGCTTCTGTAAGCGTTGTCAATGATGAGACCTTGTCAATTAACGCT 1740
QY 1741 CCAATATGACACAGAT 1800
DB 1741 CCAATATGACACAGAT 1800
QY 1801 GGCGGCAATTAAGTGAATAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1801 GGCGGCAATTAAGTGAATAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860

QY 1861 GAGGTACAGGCTTATTAACGTACCGTGGACCAAGACCTTCTGTTGGCAATTGTGAAT 1920
DB 1861 GAGGTACAGGCTTATTAACGTACCGTGGACCAAGACCTTCTGTTGGCAATTGTGAAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923
RESULT 5
AAK37279
ID AAK37279 standard; DNA; 1923 BP.
XX
XX AAK37279;
AC 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
DE Bacillus alkaline protease encoding DNA.
XX
XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KM washing composition; oxidizing agent; ss.
XX Bacillus sp.
XX MO918218-A1.
XX 15-APR-1999.
XX 07-OCT-1998; J:98MO-JP004528.
XX 07-OCT-1997; 97JP-00274570.
XX (KAOS) KAO CORP.
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hironi J, Kageyama Y;
PI Shikata S, Nomura M;
XX MPI, 1999-287736/27.
XX P-PSDB; AA17091.
XX
XX Alkali protease from Bacillus used in washing powders.
XX Disclosure; Page 63-68; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
XX
XX Sequence 1923 BP, 578 A; 417 C; 474 G; 454 T; 0 U; 0 Other;
SQ
Query Match 99.8%; Score 1919.8; DB 2; Length 1923;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGTCGACGATTTTGTGACT 60
DB 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGTCGACGATTTTGTGACT 60
QY 61 GTTGCCTTAAGTATCATCTGACGATGCTGCAAGAAATTTGATCTGATTTCAAGA 120
DB 61 GTTGCCTTAAGTATCATCTGACGATGCTGCAAGAAATTTGATCTGATTTCAAGA 120

QY 121 ATTACAGCAACAATGATGCTAAAGTTTCTCCAGCAGGGGCAAGTGGTGTGCTGCT 180
DB 121 ATTACAGCAACAATGATGCTAAAGTTTCTCCAGCAGGGGCAAGTGGTGTGCTGCT 180
QY 181 TTTCTGGTGAATCTGAAAAATGTGAACCTCCAAAGAGTTTGAGAGAAAGCTTGAACA 240
DB 181 TTTCTGGTGAATCTGAAAAATGTGAACCTCCAAAGAGTTTGAGAGAAAGCTTGAACA 240
QY 241 GTCCCGCAATATTAATCTCATATTTATCCAAATTCGAATGACCAATTTTGAAGAAACA 300
DB 241 GTCCCGCAATATTAATCTCATATTTATCCAAATTCGAATGACCAATTTTGAAGAAACA 300
QY 301 AAACAGCAGCTGCAAAAAACAGGGGCAAGATTTCTGACTACATCTGATTTATGCTTAC 360
DB 301 AAACAGCAGCTGCAAAAAACAGGGGCAAGATTTCTGACTACATCTGATTTATGCTTAC 360
QY 361 ATTGTGAGTATGAGGGCGATGTTAAGTCAGCAACAAGCACTTGAACGTGGAATCC 420
DB 361 ATTGTGAGTATGAGGGCGATGTTAAGTCAGCAACAAGCACTTGAACGTGGAATCC 420
QY 421 GTGAGCCTTATTTGCGATATACAGAAATGATCCAGCTTTTCAAAAAAGGGGCATCA 480
DB 421 GTGAGCCTTATTTGCGATATACAGAAATGATCCAGCTTTTCAAAAAAGGGGCATCA 480
QY 481 GAGCTGTAAAGCAGTGGCGCTTGATACAAAGCAAGAAATTAAGAGTGCATTTAAGA 540
DB 481 GAGCTGTAAAGCAGTGGCGCTTGATACAAAGCAAGAAATTAAGAGTGCATTTAAGA 540
QY 541 GGCATGGAACAAATGCGCAATTCGCAATPAAGCAATGATGATATTAATTAAGGGCAAG 600
DB 541 GGCATGGAACAAATGCGCAATTCGCAATPAAGCAATGATGATATTAATTAAGGGCAAG 600
QY 601 CCTGATATPAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 CCTGATATPAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 AGCAGCTACGGGTGTATGAGCAAGACAGATGTGACGGTTCGATACAGGGCTTGAT 720
DB 661 AGCAGCTACGGGTGTATGAGCAAGACAGATGTGACGGTTCGATACAGGGCTTGAT 720
QY 721 ACAGGTGCAATGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACAGGTGCAATGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCATTGGACGAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCATTGGACGAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGATTTAGAAACGGCTCCATTAATTAAGAAATGCGCTCAGGCGAATCTAGCTTTC 900
DB 841 TCCGATTTAGAAACGGCTCCATTAATTAAGAAATGCGCTCAGGCGAATCTAGCTTTC 900
QY 901 CAATCTATCATGATATGCGGTGGGGAATTTGAGAGATCTCTTCAATCTGCAAACTTCA 960
DB 901 CAATCTATCATGATATGCGGTGGGGAATTTGAGAGATCTCTTCAATCTGCAAACTTCA 960
QY 961 TTCAGGCAAGCATACGTCGTGTCGCAATTTCAATCAATCTCTGCGGAGCAAGAGTG 1020
DB 961 TTCAGGCAAGCATACGTCGTGTCGCAATTTCAATCAATCTCTGCGGAGCAAGAGTG 1020
QY 1021 AATGGGGCTTACACAACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATGGGGCTTACACAACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ACAGATCTTTTGTGCTGCGGGAATGAAGGACCAACGCGGAAACCACTGATCAACAGGC 1140
DB 1081 ACAGATCTTTTGTGCTGCGGGAATGAAGGACCAACGCGGAAACCACTGATCAACAGGC 1140
QY 1141 ACAGCTAAATAATGCAATTAACATGAGCTACGAAACCTCGCCCAAGCTTTGGGTCT 1200
DB 1141 ACAGCTAAATAATGCAATTAACATGAGCTACGAAACCTCGCCCAAGCTTTGGGTCT 1200


```

QY 1201 TATCGGAGCAATATCAACGATGAGCACTTCTCTTCACTGAGCAACGAAGATGGA 1260
DB 1201 TATCGGAGCAATATCAACGATGAGCACTTCTCTTCACTGAGCAACGAAGATGGA 1260
QY 1261 CGGATCAAAACCGGATGATGAGCAACGGGAAACGTTCACTATCAGCAAGATCTTCTTT 1320
DB 1261 CGGATCAAAACCGGATGATGAGCAACGGGAAACGTTCACTATCAGCAAGATCTTCTTT 1320
QY 1321 GCACCCGATTCCTCTCTTCTGAGGAAACATGACAGTAAATATGATATCATGGGTGGAACG 1380
DB 1321 GCACCCGATTCCTCTCTTCTGAGGAAACATGACAGTAAATATGATATCATGGGTGGAACG 1380
QY 1381 TCCATGAGCTACACGATTCGTTGCTGAGAAAGTGGACAGCTTGATGAGATTTTGAAA 1440
DB 1381 TCCATGAGCTACACGATTCGTTGCTGAGAAAGTGGACAGCTTGATGAGATTTTGAAA 1440
QY 1441 AACAGAGGATCAACACAAAGCCTTCTATTAATAAGCGGACATGATCCGGTGCAGCT 1500
DB 1441 AACAGAGGATCAACACAAAGCCTTCTATTAATAAGCGGACATGATCCGGTGCAGCT 1500
QY 1501 GACATGGGCTTGGCTTACCGGAAACGGTAACCAAGATGGGAGCAGTGAATTGATTA 1560
DB 1501 GACATGGGCTTGGCTTACCGGAAACGGTAACCAAGATGGGAGCAGTGAATTGATTA 1560
QY 1561 TCCCTGAAGCTTGCCTATGTAACGAGTCAAGTCTCTATCCACAGCCAAAGGAGAG 1620
DB 1561 TCCCTGAAGCTTGCCTATGTAACGAGTCAAGTCTCTATCCACAGCCAAAGGAGAG 1620
QY 1621 TACTCGTTTACGTCTACCTGCGGAAAGCTTTGAAAATCTCCCTGATATGCTGATGCC 1680
DB 1621 TACTCGTTTACGTCTACCTGCGGAAAGCTTTGAAAATCTCCCTGATATGCTGATGCC 1680
QY 1681 CCGGAGAGCAACCTGCTCCGTAACGCTTTCATGATGATGATGATGATGATGATGATGAT 1740
DB 1681 CCGGAGAGCAACCTGCTCCGTAACGCTTTCATGATGATGATGATGATGATGATGATGAT 1740
QY 1741 CCAAAATGAGCAACAGTATGAGAAATGACTTACTTCCCAATACATGATTAATGAGAT 1800
DB 1741 CCAAAATGAGCAACAGTATGAGAAATGACTTACTTCCCAATACATGATTAATGAGAT 1800
QY 1801 GCGCCGATTAACGTAGAAATGATTTATTAATGACCAACAAAGCGGAGCGTATACAT 1860
DB 1801 GCGCCGATTAACGTAGAAATGATTTATTAATGACCAACAAAGCGGAGCGTATACAT 1860
QY 1861 GAGGTACAGGCTTATTAAGTACCGGTTGAGCAACAGACCTTCCCTGATGATGATGAT 1920
DB 1861 GAGGTACAGGCTTATTAAGTACCGGTTGAGCAACAGACCTTCCCTGATGATGATGAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923

```

RESULT 6
AAK37277
ID AAK37277 standard; DNA; 1920 BP.

```

AC AAK37277;
XX
DT 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
DE Bacillus alkaline protease encoding DNA.
XX
KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW washing composition; oxidizing agent; ss.
XX
OS Bacillus sp.
XX
PN M09918218-Al.
XX
PD 15-APR-1999.
XX

```

```

PF 07-OCT-1998; 98MO-JP004528.
XX
XX PR 07-OCT-1997; 97JP-00274570.
XX
XX (KAOS ) KAO CORP.
XX
PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX
DR WPI; 1999-287736/27.
XX
DR P-PSDB; AAY17087, AAY17089.
XX
PT Alkali protease from Bacillus used in washing powders.
XX
XX Disclosure; Page 53-58; 71pp; Japanese.
XX
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability of washing compositions including
CC the enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 1920 BP, 576 A; 411 C; 473 G; 460 T; 0 U; 0 Other;
Query Match 76.8%; Score 1477.2; DB 2; Length 1920;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
QY 10 AAGAAAAAGGCTTTTATCTGTTTATCACTGACGAGATTTGTGACTGTTCGTTA 69
DB 7 AAGAAAGAGGCTTTTATCTGTTTATCACTGACGAGATTTGTGACTGTTCGTTA 66
QY 70 AGTAATCATCTGACGAGTGTGCAAGGAATTTGATCTGATTCGATTCGAAAGGAATTCAGCA 129
DB 67 AACATTCCTGCGCTGTGTGATGCAAGACTTTGATCTGATTCGATTCGAAAGGAATTCAGCA 126
QY 130 ACAACTAGTCTAAAGGTTTCTCCAGCAGGAGCAGCTGGTGTCTGCTTTTCTGCTG 189
DB 127 ACAACCATGTCACTGGTTTCTCCAAACAGGCAACAGCAAGTGGGCTGCACTTTCTGCTG 186
QY 190 GAATCTGAAAATGTGAAATCTCCCAAAAGCTTTCGAAAGAAAGCTTGAACAGTCCCGGCA 249
DB 187 GAGTCTGAAAATGTGAAATCTTTAAAGATGTGTAAGAAATCTGAAACAGTACCGGCA 246
QY 250 AATTAATAAATCTCATTTATTCATATTCATATGACCAATTTTGAAGAAACAAACAGCAG 309
DB 247 AATTAATAAATCTCATTTATTCATATTCATATGACCCCAATTTTGAAGAAACAAACAGCAG 306
QY 310 CTGAAAAAACAAGGAGCAAAAGATTCTGACTACATACCTGATTAATGCTTCAATTTGAG 369
DB 307 CTGAGAGCAACTGAGAGCAAAAGATTCTGACTACATCCCTGATTAATGCTTCAATTTGAG 366
QY 370 TATGAGGGGAGATGTGTAAGTCAAGCAACCAATTCAGCAGTGGAAATCCGTGAGACCT 429
DB 367 TATGAGGGGAGATGTGTAAGTCAAGCAACCAATTCAGCAGTGGAAATCCGTGAGACCT 426
QY 430 TATTTGCCGATTAACAGATATGATCCCACTTTTCAAAAGGGGATCAGAGCTTGTGA 489
DB 427 TACTTGGCGAATTAACAAATATGATCCCACTTTTCAAAAGGGGATCAGAGCTTGTG 486
QY 490 AAAGCAGTGGCGCTTGTATCAAGAGCAAAATTAAGAGGTGCAATTAAGAGCAGTGA 549
DB 487 AAAGCAGTGGCGCTTGTATCAAGAGCAAAATTAAGAGGTGCAATTAAGAGCAGTGA 546

```

QY 550 CAATCGCACAAATTCGCAATTAAGCATATGCTATATATTAAGCGAAAGCCGTAAT 609
 DB 547 GAATGCTAGTACGTAGCAAGCAATGAGTCCATTAATTAAGCGAAAGCCGTAAT 606
 QY 610 AAGGTGATGAATATGTTGCGCGTGAATTTGTCAGAACCGGATGTGCTCAGAGCAGTAC 669
 DB 607 AAGGTGATGAATATGTTGCGCGTGAATTTGTCAGAACCGGATGTGCTCAGAGCAGTAC 666
 QY 670 GGGTTGATGAGCAAGGACAGATCTGAGCGGTTGCCGATACAGGGCTTGATACAGGTGCG 729
 DB 667 GGGTTGATGAGCAAGGACAGATCTGAGCGGTTGCCGATACAGGGCTTGATACAGGGT 726
 QY 730 AATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
 DB 727 AAGGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
 QY 790 CGGACGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
 DB 787 CGGACGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
 QY 850 GGAACCGGCTCCCACTAATAAGGAAATGCGCCCTCAGGCGCAATCTAGTCTTCCATCTATC 909
 DB 847 GGAACCGGCTCCCACTAATAAGGAAATGCGCCCTCAGGCGCAATCTAGTCTTCCATCTATC 906
 QY 910 ATGATAGCGGTGGGGAATTGAGAGACTTCTGCAATCTGCAAACTTATTCAGCCAA 969
 DB 907 ATGATAGCGGTGGGGAATTGAGAGACTTCTGCAATCTGCAAACTTATTCAGCCAA 966
 QY 970 GCATACAGTCTGTGCTCCAGAAATTCATAAACTCTGCGGAGACAGGATGAAATGGGGCT 1029
 DB 967 GCATACAGTCTGTGCTCCAGAAATTCATAAACTCTGCGGAGACAGGATGAAATGGGGCT 1026
 QY 1030 TACACACAGATTCGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
 DB 1027 TACACACAGATTCGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
 QY 1090 TTCGTCGCGGGAATGAGAACCGGAAACCGGCGGAAACCATGATGACACGAGCAGATTA 1149
 DB 1087 TTCGTCGCGGGAATGAGAACCGGAAACCGGCGGAAACCATGATGACACGAGCAGATTA 1146
 QY 1150 AATGCAATTAAGTCCGAGCTACGAGAAACCTCCGCGCAAGCTTTGGGTCTTAATGCGGAC 1209
 DB 1147 AATGCAATTAAGTCCGAGCTACGAGAAACCTCCGCGCAAGCTTTGGGTCTTAATGCGGAC 1206
 QY 1210 AATATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1269
 DB 1207 AATATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
 QY 1270 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1329
 DB 1267 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
 QY 1330 TCTCTCTTCTGCGGAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1389
 DB 1327 TCTCTCTTCTGCGGAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386
 QY 1390 AACCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1449
 DB 1387 AACCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
 QY 1450 ATACAGCAAAAGCTTCTGATTAATAAGGCGGATGATGATGATGATGATGATGATGATGAT 1509
 DB 1447 ATACAGCAAAAGCTTCTGATTAATAAGGCGGATGATGATGATGATGATGATGATGATGAT 1506
 QY 1510 CTGTCGTCACCGGAAACGATTAACCAAGATGATGATGATGATGATGATGATGATGATGAT 1569
 DB 1507 CTGTCGTCACCGGAAACGATTAACCAAGATGATGATGATGATGATGATGATGATGATGAT 1566
 QY 1570 GTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1629
 DB 1567 GTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626
 QY 1630 ACTGCTAATGCGCGGCAAGCTTTGAAATATCTCCCTGATATGATGATGATGATGATGATGAT 1689

DB 1627 ACTGCAACGCGCGGCAAGCAATGAAATATCCCTGATATGATGATGATGATGATGATGATGAT 1686
 QY 1690 ACAACTGCTTCCTGTAACGCTTGTCAATGATCTGAGACCTTGATTAACGCTCCAAATGCG 1749
 DB 1687 ACTACTGCTTCCTGTAACGCTTGTCAATGATCTGAGATTTGATGATGATGATGATGATGAT 1746
 QY 1750 ACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1809
 DB 1747 ACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1806
 QY 1810 AACGTAGAAAATGATTAATTAATGACACCAAGCGGAGCTATACATTAATGATGATGATGAT 1869
 DB 1807 AACGTAGAAAATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1866
 QY 1870 GCTTATTAACGTACCGGTGAGCAACAGACCTTCTGTTGCAATGATGATGATGAT 1923
 DB 1867 GCATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920

RESULT 7
 ADY33779
 ID ADY33779 standard; DNA; 1305 BP.
 XX
 AC ADY33779;
 XX
 DT 05-MAY-2005 (first entry)
 XX
 DE Bacillus species alkaline protease coding sequence.
 XX
 KW mutagenesis; alkaline protease; detergent; laundry; cosmetic; food;
 KW bleaching detergent; gene; ds.
 OS Bacillus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1305
 FT /*tag= a
 FT /product= "alkaline protease"
 FT /note= "no start codon given"
 FT /partial
 XX
 PN EP1347044-A2.
 XX
 PD 24-SEP-2003.
 XX
 PF 21-MAR-2003; 2003BP-0006472.
 XX
 PR 22-MAR-2002; 2002JP-00081428.
 PR 06-JUN-2002; 2002JP-00165987.
 PR 18-OCT-2002; 2002JP-00304230.
 PR 18-OCT-2002; 2002JP-00304231.
 XX
 PA (KAOS) KAO CORP.
 XX
 PI Okuda M, Sato T, Saito K, Sumitomo N, Izawa Y, Saeki K;
 PI Kobayashi T, Nomura M;
 XX
 DR WPI; 2003-846540/79.
 DR P-PSDB; ADY33778.
 XX
 PT New alkaline protease having specific amino acid residue at a specific
 PT position of its amino acid sequence, useful for producing detergent
 PT compositions, laundry detergent, fiber modifiers, leather-treating agents
 PT or pipe cleaners.
 XX
 PS Disclosure; SEQ ID NO 2; 31pp; English.
 XX
 CC The invention relates to an alkaline protease having a fully defined
 CC sequence of 434 amino acids (I) given in the specification, or an amino
 CC acid sequence at least 80% homology with (I), where an amino acid residue
 CC at position 65, 101, 163, 170, 171, 273, 320, 359 or 387 of (I) is
 CC selected from 16 amino acid residues. The amino acid residues at the

corresponding positions are selected from: position 65: proline; position 101: asparagine; position 163: histidine, aspartic acid, phenylalanine, lysine, asparagine, serine, isoleucine, leucine, glutamine, threonine or valine; position 170: valine or leucine; position 171: alanine, glutamic acid, glycine or threonine; position 273: isoleucine, glycine or threonine; position 320: phenylalanine, valine, threonine, leucine, isoleucine or glycine; position 359: serine, leucine, valine, isoleucine or glutamic acid; and position 387: alanine, lysine, glutamine, glutamic acid, arginine or histidine. The alkaline protease is useful for the production of a detergent composition, such as laundry detergent, fiber conditioners, leather-treating agents, cosmetic compositions, bath additives, food modifiers and pharmaceutical compositions. The alkaline protease may also be used as bleaching detergent, hard surface cleansing detergent, pipe cleaner, artificial tooth cleaner, and as a sterilizing cleanser for medical tools. The new alkaline protease has a more potent proteolytic capacity, exhibiting excellent detergency for the removal of a complex stain, and has high secretion capacity. This sequence corresponds to the *Bacillus* sp. alkaline protease coding sequence.

Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;

Query Match 67.9%; Score 1305; DB 11; Length 1305;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 619 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAGAGCAGTACCGGTTGAT 678
DB 1 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAGAGCAGTACCGGTTGAT 60
QY 679 GGAAGAGACAGATCGTAGCGGTTGCCATACAGGGCTTGATACAGGTGCGATGACGT 738
DB 61 GGAAGAGACAGATCGTAGCGGTTGCCATACAGGGCTTGATACAGGTGCGATGACGT 120
QY 739 TCGATGCAAGAACCTTCCGCGGAAATTTATCTGCTTATATGCAATTTGGAGAGCAAT 798
DB 121 TCGATGCAAGAACCTTCCGCGGAAATTTATCTGCTTATATGCAATTTGGAGAGCAAT 180
QY 799 AATGCCAATGATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 858
DB 181 AATGCCAATGATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
QY 859 TCCATTAATTAAGAAATGAGCGGCTCAGGGAATCTAGTCTTCAATCTATCATGATGAC 918
DB 241 TCCATTAATTAAGAAATGAGCGGCTCAGGGAATCTAGTCTTCAATCTATCATGATGAC 300
QY 919 GGTGGGGGACTGAGGAGTACCTTCGATCTGCAAACTTATTCAGCCACATACAGT 978
DB 301 GGTGGGGGACTGAGGAGTACCTTCGATCTGCAAACTTATTCAGCCACATACAGT 360
QY 979 GCTGGTGCCAGATTCATCAAACTCTGGGAGAGCAGTGAATGGGCTTACACAACA 1038
DB 361 GCTGGTGCCAGATTCATCAAACTCTGGGAGAGCAGTGAATGGGCTTACACAACA 420
QY 1039 GATTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1098
DB 421 GATTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
QY 1099 GGGAAATGAAGACGAAACGGGGAACATCACTGACCAAGAGCAGCTTAAATATGATA 1158
DB 481 GGGAAATGAAGACGAAACGGGGAACATCACTGACCAAGAGCAGCTTAAATATGATA 540
QY 1159 ACAAGTCGAGCTACCGAAAACTCCGCCCAAGCTTTGGGTCTTATGCGGACAAATAC 1218
DB 541 ACAAGTCGAGCTACCGAAAACTCCGCCCAAGCTTTGGGTCTTATGCGGACAAATAC 600
QY 1219 CATGTGGACAGCTTCTTCACTGAGTGAACCAAGATGAGCGGATCAAAACCGATGTC 1278
DB 601 CATGTGGACAGCTTCTTCACTGAGTGAACCAAGATGAGCGGATCAAAACCGATGTC 660
QY 1279 ATGGACCGGGAAAGTTTCACTATCAGAGAGATCTTCTTGGACCGGATTCCTCTTC 1338
DB 661 ATGGACCGGGAAAGTTTCACTATCAGAGAGATCTTCTTGGACCGGATTCCTCTTC 720

```

```

QY 1339 TGGGGAAACATGACAGTAAATATGATATGAGTGGAGAACTCCATGCTTACACCGATC 1398
DB 721 TGGGGAAACATGACAGTAAATATGATATGAGTGGAGAACTCCATGCTTACACCGATC 780
QY 1399 GTTGTGGAACAGTGGACAGAGCTTGGTGAAGATTTTGTGAAGAAACAGAGGATCAACCA 1458
DB 781 GTTGTGGAACAGTGGACAGAGCTTGGTGAAGATTTTGTGAAGAAACAGAGGATCAACCA 840
QY 1459 AAGCTTCTCTATTAAGACGCACTGATTTGCGGTGACGCTGATCGGCTTGGCTAC 1518
DB 841 AAGCTTCTCTATTAAGACGCACTGATTTGCGGTGACGCTGATCGGCTTGGCTAC 900
QY 1519 CCGAAGCGTAACCAAGATGGGAGAGTGAATGATGATGATGATGATGATGATGATGATG 1578
DB 901 CCGAAGCGTAACCAAGATGGGAGAGTGAATGATGATGATGATGATGATGATGATGATG 960
QY 1579 GTGAACGAGTCAAGTCTCTATTCACACGACCAAGGAGAGAGTACCTGTTACTGCTACT 1638
DB 961 GTGAACGAGTCAAGTCTCTATTCACACGACCAAGGAGAGAGTACCTGTTACTGCTACT 1020
QY 1639 GCCGCAAGCTTTGAAATCTCCCTGTGATGATGATGATGATGATGATGATGATGATG 1698
DB 1021 GCCGCAAGCTTTGAAATCTCCCTGTGATGATGATGATGATGATGATGATGATGATG 1080
QY 1699 TCCGTAAGCTTGTGAATGATCTGGAACCTTGTCAATTAACGCTCAAAATGGACACAGAT 1758
DB 1081 TCCGTAAGCTTGTGAATGATCTGGAACCTTGTCAATTAACGCTCAAAATGGACACAGAT 1140
QY 1759 GTAGAAATGACCTTACTCTGCAATCAATGATGATGATGATGATGATGATGATGATG 1818
DB 1141 GTAGAAATGACCTTACTCTGCAATCAATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1819 AATGATTTTATTAATGACCAACAAACGGGAGCTTATCAATTAAGTACAGGCTTATTAAC 1878
DB 1201 AATGATTTTATTAATGACCAACAAACGGGAGCTTATCAATTAAGTACAGGCTTATTAAC 1260
QY 1879 GTACCGGTTGACCAACAGACCTTCTCGTTGGCAATTTGATTTAA 1923
DB 1261 GTACCGGTTGACCAACAGACCTTCTCGTTGGCAATTTGATTTAA 1305

RESULT 8
ADL25803
ID ADL25803 standard; DNA; 1305 BP.
XX
AC ADL25803;
XX
DT 20-MAY-2004 (first entry)
XX
DE Bacillus alkaline protease coding sequence.
XX
KM alkaline protease; washing agent; gene; ds; enzyme.
XX
OS Bacillus sp.
XX
FH Key
XX CDS 1..1305
XX FT location/Qualifiers
XX FT /*tag= a
XX FT /partial
XX FT /product= "Bacillus alkaline protease"
XX FT /note= "No start codon is given"
XX
XX JP2004008085-A.
XX
XX 15-JAN-2004.
XX
XX 06-JUN-2002; 2002JP-00165950.
XX
XX 06-JUN-2002; 2002JP-00165950.
XX
XX (KAOS ) KAO CORP.
XX
XX WPI; 2004-094297/10.

```

DR P-PSDB; ADL25802.

XX Novel mutant alkaline protease produced by substituting the amino acid
PT residues useful as washing agent.

XX Disclosure; SEQ ID NO 2; 21bp; Japanese.

XX The invention comprises a mutant *Bacillus* alkaline protease which
CC contains substitutions at positions 163, 170 and 434. The mutant alkaline
CC protease is useful as a washing agent. The present DNA sequence encodes a
CC *Bacillus* alkaline protease of the invention.

XX Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;

Query Match 67.9%; Score 1305; DB 12; Length 1305;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 619 AATGATGTTGCGCGTGAATTTGCAAGCGGATGAGCTGAGAGCAAGCTAGGGTTGTAT 678
DB 1 AATGATGTTGCGCGTGAATTTGCAAGCGGATGAGCTGAGAGCAAGCTAGGGTTGTAT 60
QY 679 GGAAGAAGACAGATCGTAGCGGTTGCCGATACAGGGCTTGATACAGTTCGCAATGACAGT 738
DB 61 GGAAGAAGACAGATCGTAGCGGTTGCCGATACAGGGCTTGATACAGTTCGCAATGACAGT 120
QY 739 TCGATGCAATGAAGCTTTCCGCGGAAAATTACTGCAATATATGCAATTTGGAGCGAGAT 798
DB 121 TCGATGCAATGAAGCTTTCCGCGGAAAATTACTGCAATATATGCAATTTGGAGCGAGAT 180
QY 799 AATGCCAATGATACGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 858
DB 181 AATGCCAATGATACGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 859 TCCACTAATTAAGAAATGCGGCTCAGCGCAATCTAATCTTCCATCTATCATGATAGC 918
DB 241 TCCACTAATTAAGAAATGCGGCTCAGCGCAATCTAATCTTCCATCTATCATGATAGC 300
QY 919 GGTGGGGGACTTGGAGAGTACTCTTGATCTGCAAACTTTATTCAGCCAGATACAGT 978
DB 301 GGTGGGGGACTTGGAGAGTACTCTTGATCTGCAAACTTTATTCAGCCAGATACAGT 360
QY 979 GGTGGTCCGAGAATTCATCAAACTCTGGGAGAGCAGTGAATGGGCTTACACAACA 1038
DB 361 GGTGGTCCGAGAATTCATCAAACTCTGGGAGAGCAGTGAATGGGCTTACACAACA 420
QY 1039 GATTCCAGAAATGTGATGATCTATGTGCGCAAAATGATATGATCTCTTTGCGTGGC 1098
DB 421 GATTCCAGAAATGTGATGATCTATGTGCGCAAAATGATATGATCTCTTTGCGTGGC 480
QY 1099 GGGAAATGAAGACCGAAGCGGGAACATCATGTCACAGGCAAGCTTAAATGCAAT 1158
DB 481 GGGAAATGAAGACCGAAGCGGGAACATCATGTCACAGGCAAGCTTAAATGCAAT 540
QY 1159 ACAAGTCGAGCTACCGGAAAACCTCGCGCAAGCTTTGGGCTTATAGCGGCAATATCAAC 1218
DB 541 ACAAGTCGAGCTACCGGAAAACCTCGCGCAAGCTTTGGGCTTATAGCGGCAATATCAAC 600
QY 1219 CATGTGCGCAGATTCTTCAAGTGAACCGAACAAGATGAGCGATCAAAACCGAGATGC 1278
DB 601 CATGTGCGCAGATTCTTCAAGTGAACCGAACAAGATGAGCGATCAAAACCGAGATGC 660
QY 1279 AATGGACCGGGAAAGTTCTACTATATGACGAAGATCTTCTTGACCGGATTCCTCTTC 1338
DB 661 AATGGACCGGGAAAGTTCTACTATATGACGAAGATCTTCTTGACCGGATTCCTCTTC 720
QY 1339 TGGGCGAAGCATGACGATAATATGATCATGCGGGGAACGTCATGCTACACGAGATC 1398
DB 721 TGGGCGAAGCATGACGATAATATGATCATGCGGGGAACGTCATGCTACACGAGATC 780
QY 1399 GTTGTCTGAAAAGTGTGACAGCTTCTGTGACATTTTGTGAAAAACAGAGCATCACCA 1458
DB 781 GTTGTCTGAAAAGTGTGACAGCTTCTGTGACATTTTGTGAAAAACAGAGCATCACCA 840

```

```

QY 1459 AAGCTTCTCTATTAAGGCGACATGATTCGCGGTCAGCTGACATCGGCTTGGCTAC 1518
DB 841 AAGCTTCTCTATTAAGGCGACATGATTCGCGGTCAGCTGACATCGGCTTGGCTAC 900
QY 1519 CCGAAGCTTAAACGAAGATGGGAGAGTGAATGATGATTAATCCCTGAACGTTGCTAT 1578
DB 901 CCGAAGCTTAAACGAAGATGGGAGAGTGAATGATGATTAATCCCTGAACGTTGCTAT 960
QY 1579 GTGAACGAGTCCAGTTCTCTATTCACCAAGCAAAAGCGACGTACTGTTACTGCTACT 1638
DB 961 GTGAACGAGTCCAGTTCTCTATTCACCAAGCAAAAGCGACGTACTGTTACTGCTACT 1020
QY 1639 GCCGGAAGCTTTGAAAATCTCCCTGATATGATGATGATGATGATGATGATGATGAT 1698
DB 1021 GCCGGAAGCTTTGAAAATCTCCCTGATATGATGATGATGATGATGATGATGATGAT 1080
QY 1699 TCCGTAACGCTTGTCAATGATCTGACCTTGTCAATACCGCTCAATGAGCACAGTAT 1758
DB 1081 TCCGTAACGCTTGTCAATGATCTGACCTTGTCAATACCGCTCAATGAGCACAGTAT 1140
QY 1759 GTAGAAATGACTTTACTTCCGCAATCAATGATGATGATGATGATGATGATGATGAT 1818
DB 1141 GTAGAAATGACTTTACTTCCGCAATCAATGATGATGATGATGATGATGATGATGAT 1200
QY 1819 AATGATTTTATTAATGACCAACAAGCGGAGCTATCAATGAGTACAGGCTTATAC 1878
DB 1201 AATGATTTTATTAATGACCAACAAGCGGAGCTATCAATGAGTACAGGCTTATAC 1260
QY 1879 GTACGGGTTGAGCAAGACGCTTCTGTTGGCAATTGTAATTAA 1923
DB 1261 GTACGGGTTGAGCAAGACGCTTCTGTTGGCAATTGTAATTAA 1305

```

```

RESULT 9
AAV82382
ID AAV82382 standard; DNA; 3003 BP.
XX
AC AAV82382;
XX
DT 12-APR-1999 (first entry)
XX
DE Bacillus Jp170 protease gene.
XX
KW Protease; detergent; surfactant; leather processing; deblistering;
XX flavour; ss.
XX
OS Bacillus sp.
XX
FH Key Location/Qualifiers
FT CDS 846..2771
FT FT /*tag= a
FT sig_peptide 846..944
FT FT /*tag= b
FT mat_peptide 1470..2768
FT FT /*tag= c
XX
XX WO9856927-A2.
XX
XX 17-DEC-1998.
XX
XX 09-JUN-1998; 98WO-US012005.
XX
XX 12-JUN-1997; 97US-00873479.
XX
XX (NOVO ) NOVO NORDISK BIOTECH INC.
XX
XX Sloma A, Christlanon L;
XX
XX WPI; 1999-080908/07.
XX
XX P-PSDB; AAV89547.
XX
XX Novel protease from Bacillus subtilis LC20 - useful in laundry and

```

dishwashing detergents and for leather processing.
 Claim 11: Page 52-53; 77pp; English.

This nucleotide sequence encodes a novel protease (see AAW82382) of *Bacillus* sp. JP170 (NCIB 12513). The sequence in plasmid p170BAN is contained in *Bacillus subtilis* LC20 NRRL B-21680. The protease gene was isolated from chromosomal DNA of JP170 following preparation of probes based on protease N-terminal and internal peptides (see AAW8549-50), screening of chromosomal libraries, isolation of the 3' end of the gene by inverse PCR (see AAW82410-11), reconstruction of 5' and 3' ends and PCR amplification (see AAW82412-16). Claimed recombinant host cells can be used in a method for producing the protease. The protease is used in laundry and dishwashing detergents, for institutional and industrial cleaning, and for leather processing, as well as for debittering and enhancing the degree of hydrolysis of protein hydrolysates, for flavour development through hydrolysis of proteins, degradation of undesired peptides and in enzymatic synthesis of peptides. It has enhanced stability towards oxidation under alkaline conditions, e.g. towards bleaching agents of the peroxy type. The invention also provides mutant cells in which the protease activity is diminished. Such cells can be used for the production of heterologous recombinant proteins.

Sequence 3003 BP; 982 A; 504 C; 645 G; 872 T; 0 U; 0 Other;
 Query Match 100%; Score 1166.4; DB 2; Length 3003;
 Best Local Similarity 76.0%; Pred. No. 0;
 Matches 1453; Conservative 0; Mismatches 456; Indels 3; Gaps 1;

12 GAAAAAGTGTGTTTATCTGTTTATCAGCTGACGAGTATTTGTCGTCGTTAAG 71
 863 GAAGAGGGTGTGTTTATCCGTTTATCAGTTCCTCACAATATGTCCTGTTGCTTAA 922
 72 TATATCATCTGACGGTGTGCAAGAAATTTGATCTGATTTCAAGAAATCAACAC 131
 923 CAGTCTTCTTACTATTTGGGCGCAACATTTGAATTTGACCTTTAAGGGGATGAACACT 982
 132 AACTGATGTAAAGGTTTCTCCAGCAGGGGCGACGTCGTGCTGCTTTCTGATGGA 191
 963 TACGCTAAGAAAGGCTGCCACACACAGAAAGAAAGGAAAGGATCTTTCTGTAA 1042
 192 ATCTGAAATGTTGAAATCTCCAAAGGTTTGGCAAGAAAGCTTGAACAGTCCGGCA 251
 1043 CTCTGAAATGTTGAAATCTCCAAAGGTTTGAAGAAAGCTTGAAGAGTCCAGCGGA 1102
 252 TATATACTCATATATTCATTCATTCATGACCAATTTTGAAGAAAGAAACAGCACT 311
 1103 TACCAAGCTATATATCTTCAATTTGACGACCTATTTTGAAGAAAGCAACCTCAACT 1162
 312 GGAAGAAACAGGGGCAAGATTTCTGACTACATACCTGATTTATGCTTACATTTGCA 371
 1163 AGAGAAAGCGGGAGCAAAATTTCTGATTAACATACAGATTTACGTTATTTGTCGAT 1222
 372 TGAAGGCGCATGTTAAGTCAAGCAAGCAACCTTGAAGCACTGTAATCCGTGAGACCT 431
 1223 TGAATGGGATGTTAAAGGCGGTAACTTACGCAATTCGCAATTTGGATGGTGAACATA 1282
 432 TTGGCGATATACAGATAGATCCCACTTTTCAAAAAGGGGCGATCAGAGCTTGTAA 491
 1283 TTTACCTTTATATATATATGACCGCACTTATTTTCCAGAGAGCTTTGAAATTAATGA 1342
 492 AGCAGTGGCGTTGATCAAGCAAGCAAAATTAAGAGTGCATTTAAGAGCATCGACA 551
 1343 AACAGTGGCTTTAGTAAAGAAAGCAAGAAAGTAAAGTAAAGCTTTAAGAGATTTGA 1402
 552 AATGCAACAATTCGCAATTAAGCAATGATGCTATATTTACGGGAAAGCTTGATTA 611
 1403 AATGCCCCAATAACCGCAAAATTAATGATGTTATTAACGTAACCCCAAGCCCTGAAT 1462
 612 GGTGATGATGATGTTGGCGCTGGAATTTGCAAAAGCGGATGCTGAGACAGTACCG 671
 1463 AGTTTGAATGACGTGGCCGTGCTGATGTAAGAGAGAGCTGCGCAAAATTAATCTTTGG 1522

672 GTTGTATGACAAAGACAGATGCTGACGGTTCGCAATCAGGGCTTGAATACAGTTCGCA 731
 1523 CTATATATGACAAAGACAGATGCTGACGGTTCGCAATCAGGGCTTGAATACAGGAA 1582
 732 TGAACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 791
 1583 TGAACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1642
 792 GACGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 851
 1643 AACGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1702
 852 AAGCGCTTCAATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 911
 1703 AATG--CTCAAAATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1759
 912 GATATGCGTGGGAGCTTGGAGGACTCTGCAATCTGCAAACTTATTCAGGCAAGC 971
 1760 GATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1819
 972 ATACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1031
 1820 ATATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1091
 1032 CACAAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1151
 1880 TACAGCAGATCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1199
 1092 CGTTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1211
 1940 TCGCGCGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1271
 2000 TGCATTTACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1282
 1212 TATCAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1331
 2060 TATTTACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1399
 1272 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1451
 2120 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1511
 1332 CTCTTCTGCGGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1571
 2180 CTCAATCTGCGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1631
 1392 ACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1691
 2240 TCGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1751
 1452 CACACAAAGCTTCTATTTAAAGGCGCATGATGATGATGATGATGATGATGATGATG 1811
 2300 AACTCTTAAGCTTCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTT 1871
 1512 TGGTACCGGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1931
 2360 TGGCTTTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1991
 1572 TGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2051
 2420 CGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2111
 1632 TGCATCTGCGGCAAGCTTGAATATGATGATGATGATGATGATGATGATGATGATGATG 2171
 2480 GGTCTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2231
 1692 AACTGCTTCAAGCTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATG 2291
 2540 GAGCGATCTCAATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2351
 1752 ACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2411

Db 1258 GTACCATCTGGCCACAGCGTTTCTGACTAGTATGTAACAT 1299

RESULT 11

ADA71938
ID ADA71938 standard; DNA; 2000 BP.

XX
XX ADA71938;

XX
XX 20-NOV-2003 (first entry)

XX
XX DE Rice gene, SEQ ID 5263.

XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds. 11.

XX
XX Oryza sativa.

XX
XX WO2003000898-A1.

XX
XX PD 03-JAN-2003.

XX
XX 22-JUN-2001; 2001WO-1B001105.

XX
XX PR 22-JUN-2001; 2001WO-1B001105.

XX
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX
XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
PI Katagiri F, Qian S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;
DR WPI; 2003-175290/17.

XX
XX PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

XX
XX PS Claim 27; SEQ ID NO 5263; 899pp; English.

XX
XX CC The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance to a plant to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.

XX
XX SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

XX
XX Query Match 11. 2.8%; Score 53.6; DB 8; Length 2000;
XX Best Local Similarity 8.4%; Pred. No. 8.8e-05;
XX Matches 62; Conservative 342; Mismatches 336; Indels 2; Gaps 1;

Qy 355 GCTTACATGTCGATGATGAGGGCGATGTAATGACGAACAAGACCATGAGAGGTG 414
Db 213 RYGTSRKTYGTYKKTYYASRCRAHYTTTSYNAACSSYTCRSKRSMWMMKRRKRWMS 272
Qy 415 GAATCCGTGAGCCTTATTTGCGCATATACAAATAGATCCCGACCTTTGACAAAAGG 474
Db 273 RSYGMYSWYMMWCTAYKKSYSRMYRGGRGATRYMGVSRMAMMYKMYWY 332
Qy 475 GCATCAGAGCTTGTAAGACAGTGGCGCTGATACAAAGCAAAAATAAGAGGTGCA 534
Db 333 RGYGMKRGWAGSRMMSRCWSPACYMYRMWRMRTRRRRMAKKSRTSRKRKRKCM 392
Qy 535 TTAAGAGCATGCAACAATGCAATTCGCAATTAAGCAATGATGTCATATATTAACG 594
Db 393 RKRKYKMBRGYSRRSRSCBAAWMBRCRSGRAMWRCGCGMTCRKYSIGMRMRKSWKMAAS 452

Qy 595 GCAAAAGCCTGATATTAAGTGATGATGATGTTGCGCGTGAATTCGAAAGCGAT--G 652
Db 453 KYKMSRPRYRRKCKKCSRTTYMGKTRGGMGMGTGRCRYKRSGMKRCRRRRMGMYMR 512
Qy 653 TGGCTCAGAGCAGTACCGGTTGTATGACAAGACAGATGCTGCGGTTCCGATACAG 712
Db 513 MKRYMSARPYMYRYARAKKYSYNAARKRCYRGHYTMAAGMMKRYRDMYTKMYMYK 572
Qy 713 GCGTTGATACAGTGGCAATGACAGTTGATGATGATGATGATGATGATGATGATGATG 772
Db 573 RYKSKSWYCKMSYASCMKARKAGACMKRSKMSASKMSRSRKRKXASRSSAKR 632
Qy 773 CATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 832
Db 633 YAMMGATSGSRMSBWSYTCYMRWGSMSKCTMYMYMSKYTAKGSYRYRYRMCM 692
Qy 833 TGGCTGCTCGCTATTAGAAACGCTCCATTAATAAGAAATGCGCGCTGACGAGATC 892
Db 693 YMMRMYRYRYSYTYMAWYTSSTRMAMTGKYSGRYTWTSYKCKSKSRSMYTYSW 752
Qy 893 TAGTCTTCAATCTATCATGATGATGATGATGATGATGATGATGATGATGATGATG 952
Db 753 WMAKTKWMMRBYATRMWMMYRYSMKYTYWCTMGYWWMYRYMYKRYMYKCTKYTYW 812
Qy 953 AAACCTTATGAGCGCAACATACAGTCTGCTGCGCAATTCATCAAACTCTGGGAG 1012
Db 813 WSAITYTGTWAMWMAKTKWMMGTAKTRGRARARAYMMWMAWTCAKMMTKGAKAKW 872
Qy 1013 CAGCAGTGAATGGGCTTACACACAGATTCACAGAAATGATGATGATGATGATG 1072
Db 873 AWTWMAKAMRYKYSWMBRAWYIYKTRIRTKTCMMARBSWMAWMMWMAKMAWMMW 932
Qy 1073 ATGATATGACATCTTTTGGC 1094
Db 933 KGRMGWTKYMYWCTTWMAAC 954

RESULT 12

ADA71938/C

ID ADA71938 standard; DNA; 2000 BP.

XX
XX AC ADA71938;

XX
XX AC 20-NOV-2003 (first entry)

XX
XX DE Rice gene, SEQ ID 5263.

XX
XX KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.

XX
XX OS Oryza sativa.

XX
XX WO2003000898-A1.

XX
XX PD 03-JAN-2003.

XX
XX PE 22-JUN-2001; 2001WO-1B001105.

XX
XX PR 22-JUN-2001; 2001WO-1B001105.

XX
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX
XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
PI Katagiri F, Qian S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;
DR WPI; 2003-175290/17.

XX
XX PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

XX
XX PS Claim 27; SEQ ID NO 5263; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding resistance or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

XX Query Match 2.6%; Score 50; DB 8; Length 2000;
 XX Best Local Similarity 8.0%; Pred. No. 0.0011;
 XX Matches 41; Conservative 243; Mismatches 228; Indels 0; Gaps 0;

QY 517 AAAAAAAGAGGTGCAATTAAGAGCATGCAACAATTCGCAATTCAGCAAT 576
 DB 949 AAGRRWRRAWKCVCCKMKMKMTSCMKMTKMTSCMTTMMGARRATYARRRRRTY 890
 QY 577 GATGTCATATATATTAACGCAAGCCTGATATAGTGATGATGTCGCGTGA 636
 DB 889 KMSWRMYTMTKMAWTMTTMCMAKMYATGATWMMRYTYCYAMTCKKCYKMA 830
 QY 637 ATTGTCAAGCGGATGCTGCTGAGACAGCTACGGGTTGATGACAAGACATCTGA 696
 DB 829 MTKMWTYMACWATYTSWRRAWAGMKYKMRATYWWMRMCWAKGAMRMKSRMYRM 770
 QY 697 GGGGTCGCGATACAGGCTGATGATGACGATGACATGATGATGATGACGCTTC 756
 DB 769 KATATRYKMMMTMTTMMWSRRKSYRMSGMRMWSAMRYCSRMKCACTKASSARMT 710
 QY 757 CGCGGAAATTAATTCATATATATGATTGGACGACGCAATATATCCATGATACGAAT 816
 DB 709 KKAKSRYRBRMYRMRKMGTYRYSRMBRABMSRRKAGASCMWYMGAS 650
 QY 817 GGTGATGTAACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 876
 DB 649 MMYSKTSCSAKCKKRYMTSSYMSYMSYMSYMSYMSYMSYMSYMSYMSYMSYMSY 590
 QY 877 GGGCCTCAGCGAATGATGCTTCAATGATGATGATGATGATGATGATGATGATGATG 936
 DB 589 STRSRMSGWSGSRMYTMRMKMRKRYTMRKCTYMRKCTYMRKCTYMRKCTYMRKCTY 530
 QY 937 CTACCTTCGAATCTGCAAACTTATTCAGCAAGCATAGCTGCTGCTGCTGCTGCTGCT 996
 DB 529 YTGRTYARTYSRRMYMYRKYCWYYTGYMTKCSYMYRXYGCTKCKCTAMCKKAY 470
 QY 997 ACAAACTCCTGGGAGACAGCACTGAATGGGCG 1028
 DB 469 GCMWYMYRKYSKMMSSTKYMMSWYKCKRS 438

RESULT 13
 AD333797
 ID ADY33797 standard; DNA; 54 BP.

XX ADV33797;
 XX
 DT 05-MAY-2005 (first entry)
 XX
 DE Bacillus species alkaline protease coding sequence mutagenic primer #18.
 XX
 KW mutagenesis; alkaline protease; detergent; laundry; cosmetic; food;
 XX bleaching detergent; primer; PCR; ss.
 OS Bacillus sp.
 OS Synthetic.
 XX
 PN EP1347044-A2.
 XX

PD 24-SEP-2003.

XX 21-MAR-2003; 2003EP-00006472.
 XX
 XX 22-MAR-2002; 2002JP-00081428.
 PR 06-JUN-2002; 2002JP-00165987.
 PR 18-OCT-2002; 2002JP-00304230.
 PR 18-OCT-2002; 2002JP-00304231.
 XX
 XX (KAO) KAO CORP.

PI Okuda M, Sato T, Saito K, Sumitomo N, Izawa Y, Saeki K;
 PI Kobayashi T, Nomura M;
 DR WPI; 2003-846540/79.
 XX

PT New alkaline protease having specific amino acid residue at a specific
 PT position of its amino acid sequence, useful for producing detergent
 PT compositions, laundry detergent, fiber modifiers, leather-treating agents
 PT or pipe cleaners.

XX Example 3; SEQ ID NO 20; 31pp; English.

XX The invention relates to an alkaline protease having a fully defined
 CC sequence of 434 amino acids (I) given in the specification, or an amino
 CC acid sequence at least 80% homology with (I), where an amino acid residue
 CC at position 65, 101, 163, 170, 171, 273, 320, 359 or 387 of (I) is
 CC selected from 16 amino acid residues. The amino acid residues at the
 CC corresponding positions are selected from: position 65; proline; position
 CC 101; asparagine; position 163; histidine; aspartic acid; phenylalanine;
 CC lysine; asparagine; serine; isoleucine; leucine; glutamine; threonine or
 CC valine; position 170; valine or leucine; position 171; alanine, glutamic
 CC acid, glycine or threonine; position 273; isoleucine, glycine or
 CC threonine; position 320; phenylalanine, valine, threonine, leucine,
 CC isoleucine or glycine; position 359; serine, leucine, valine, glutamine
 CC or glutamic acid; and position 387; alanine, lysine, glutamine, glutamic
 CC acid, arginine or histidine. The alkaline protease is useful for the
 CC production of a detergent composition, such as laundry detergent, fiber
 CC modifiers, leather-treating agents, cosmetic compositions, bath
 CC additives, food modifiers and pharmaceutical compositions. The alkaline
 CC protease may also be used as bleaching detergent, hard surface cleansing
 CC detergent, pipe cleaner, artificial tooth cleaner, and as a sterilizing
 CC cleanser for medical tools. The new alkaline protease has a more potent
 CC proteolytic capacity, exhibiting excellent detergent for the removal of
 CC a complex stain, and has high secretion capacity. This sequence
 CC corresponds to a PCR primer to mutagenize the Bacillus sp. alkaline
 CC protease coding sequence of the invention.

XX Sequence 54 BP; 12 A; 17 C; 16 G; 3 T; 0 U; 6 Other;

XX Query Match 2.5%; Score 48; DB 11; Length 54;
 XX Best Local Similarity 88.9%; Pred. No. 0.0007;
 XX Matches 48; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1092 CGTCGCGGGAATGAGACCGAAGCGGGAACCATCACTGCAACGACGACGCG 1145
 DB 1 CGTCGCGGGAATNNNGAACCGAAGCGGGAACCATCANNNGACCAAGGACGACG 54

RESULT 14
 ADL25809
 ID ADL25809 standard; DNA; 54 BP.

XX ADL25809;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Bacillus alkaline protease mutagenic PCR primer #4.
 XX
 KW alkaline protease; washing agent; mutagenic PCR; ss; primer.
 XX
 OS Bacillus sp.
 OS Synthetic.
 XX

XX JP2004008085-A.
 PN 15-JUN-2004.
 PD 06-JUN-2002; 2002JP-00165950.
 PF 06-JUN-2002; 2002JP-00165950.
 PR (KAOS) KAO CORP.
 PA WPI; 2004-094297/10.
 DR Novel mutant alkaline protease produced by substituting the amino acid
 PT residues useful as washing agent.
 PT Example 1; SEQ ID NO 8; 21pp; Japanese.
 XX The invention comprises a mutant Bacillus alkaline protease which
 CC contains substitutions at positions 163, 170 and 434. The mutant alkaline
 CC protease is useful as a washing agent. The present DNA sequence
 CC represents a mutagenic PCR primer for the Bacillus alkaline protease
 CC gene.
 XX Sequence 54 BP; 12 A; 17 C; 16 G; 3 T; 0 U; 6 Other;
 SQ
 Query Match 2.5%; Score 48; DB 12; Length 54;
 Best Local Similarity 88.9%; Pred. No. 0.0007; Mismatches 6; Indels 0; Gaps 0;
 Matches 48; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1092 CGCTCCCGGAAATGAAGACCGACGCGAATCATCATGTCACGACGACGAC 1145
 DB 1 CGCTCCCGGAAATNNNGACCGACGCGGAAATCATCATNNNGACGACGACGAC 54
 RESULT 15
 AB082958/C
 ID AB082958 standard; cDNA; 415 BP.
 AC AB082958;
 XX 15-JAN-2003 (first entry)
 DT Human lung specific nucleic acid (LSNA) SEQ ID NO:101.
 DE Human; lung specific gene; lung specific nucleic acid; LSG; LSNA; LSP;
 XX Lung specific protein; lung cancer; vaccine; gene therapy; gene; ss.
 KW Homo sapiens.
 XX WO200277236-A2.
 PN 03-OCT-2002.
 PD 26-OCT-2001; 2001WO-US050385.
 PP 26-OCT-2000; 2000US-0243459P.
 PR (DIAD-) DIADEXUS INC.
 PA Recipon H, Sun Y, Chen S, Liu C, Turner LR;
 PI WPI; 2003-018930/01.
 DR New isolated nucleic acid molecule, useful for treating lung cancer, and
 PT diagnosing or monitoring the presence of metastases of lung cancer in a
 PT patient.
 XX Claim 1; Page 201; 262pp; English.
 PS The present invention describes a isolated human lung specific nucleic
 CC acid (LSNA) molecule comprising a sequence selected from AB082858 to
 CC AB082993, which encodes a human lung specific protein (LSP) sequence

CC selected from ABP54165 to ABP54272. Also described: (1) a method for
 CC determining the presence of a LSNA or LSP in a sample; (2) a vector
 CC comprising the novel nucleic acid; (3) a host cell comprising the vector
 CC of (2); (4) producing a polypeptide encoded by the novel nucleic acid;
 CC (5) an isolated polypeptide produced by the method of (4); (6) an
 CC antibody or its fragment that specifically binds to the polypeptide of
 CC (5); (7) diagnosing and monitoring the presence of metastases of lung
 CC cancer in a patient; (8) treating a patient with lung cancer; or (9) a
 CC vaccine comprising the polypeptide of (5) or the novel nucleic acid. The
 CC human LSNA and LSP sequences have cytosstatic activity, and can be used in
 CC for treating lung cancer. The LSNA and LSP sequences can also be used in
 CC metastases of lung cancer in a patient
 XX
 SQ Sequence 415 BP; 117 A; 64 C; 79 G; 155 T; 0 U; 0 Other;
 Query Match 2.3%; Score 45; DB 8; Length 415;
 Best Local Similarity 60.0%; Pred. No. 0.016; Mismatches 50; Indels 0; Gaps 0;
 Matches 75; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 QY 486 TGTAAAGCAGTGCCTTGTATGAAAGCAAAATTAAGAGTGCATTAAGAGCAT 545
 DB 255 TGAATAAGCAATCCCAATTGATAGCAAGTAAATATCAATAGAGAAATTAATAGCAA 196
 QY 546 CGAACAAATGCGACAAATTCGCAATAGCAATGATGCTATATTAACGCAAGCCTGA 605
 DB 195 AGATTAAGCATCTCACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 136
 QY 606 GTATA 610
 DB 135 GAAAA 131
 Search completed: April 8, 2006, 03:01:43
 Job time : 2991.33 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 01:05:15 ; Search time 6945.92 Seconds
(without alignments)
12953.156 Million cell updates/sec

Title: US-10-784-870-5

Perfect score: 1923

Sequence: 1 atgagagaagaagaaaggt.....cgttgcaattgtgattaa 1923

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the change being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	44.8	2.3	1101	10	CNS0039G
2	43	2.2	486	5	BU682328 UI-CF-BC1
3	42.4	2.2	500	1	AI268988 gK04d01.x
4	42.4	2.2	1201	10	CNS0102T
5	42	2.2	633	3	BU369190
6	42	2.2	631	3	BU393925
7	41.8	2.2	457	1	AI689929 tX31909.x
8	41.8	2.2	464	1	AA399583 tX31904.s
9	41.8	2.2	522	1	AI961026 wG58d01.x
10	41.8	2.2	561	1	AM572909 hF17d03.x
11	41.8	2.2	628	6	CM440627 UI-H-SD0-
12	41.8	2.2	678	10	CM901981
13	41.8	2.2	685	10	CM901981
14	41.8	2.2	7301	4	HSK805591
15	41.4	2.2	449	3	BSJ59939
16	41.4	2.2	997	10	CNS005TB
17	41.2	2.1	750	5	BQ743928
18	41.2	2.1	770	9	BZ069516
19	41.2	2.1	914	8	DN483960
20	41	2.1	440	1	AA865302
21	41	2.1	482	7	CR743947
22	41	2.1	486	5	BX112112

23	41	2.1	519	1	AI732955
24	41	2.1	531	1	AI762342
C 25	41	2.1	983	10	CLS14519
26	40.8	2.1	625	7	CV250509
27	40.8	2.1	1101	10	CNS00CNT
28	40.6	2.1	538	1	AA037767
29	40.6	2.1	648	8	DR063487
C 30	40.6	2.1	1253	10	AG390521
C 31	40.4	2.1	341	7	CK479039
C 32	40.4	2.1	659	10	CM154791
C 33	40.4	2.1	1101	10	CNS00240
C 34	40.2	2.1	854	6	CB314560
C 35	40	2.1	433	1	AU300716
C 36	40	2.1	523	7	CM487402
C 37	40	2.1	575	10	CM856704
C 38	40	2.1	588	1	AU260495
C 39	40	2.1	764	7	CJ361543
C 40	40	2.1	766	7	CJ345334
C 41	40	2.1	944	7	CJ363875
C 42	39.8	2.1	484	3	BU334651
C 43	39.8	2.1	601	3	BU387574
C 44	39.8	2.1	674	5	BQ504480
C 45	39.8	2.1	828	1	AL706691

ALIGNMENTS

RESULT 1
CNS0039G/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TBT3 end of BAC # BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL063921
VERSION
AL063921.1 GI:4941778
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 1101)
AUTHORS
Direct Substation
TITL
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
JOURNAL
- Web : www.genoscope.cns.fr)

COMMENT
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Oono and Aaron Mammeter in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain v2; cn bw ap, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08K10"
/clone_1b="RPCI-98"
/note="end : TBT3"

ORIGIN

[illegible]

RESULT 2			
BUE82328			
LOCUS			
DEFINITION	BUE82328	486 bp	mRNA
ACCESSION	U1-CF-BC1-acb-g-17-0-U1.s1	U1-CF-BC1 Homo sapiens	linear EST 07-OCT-2002
VERSION	BUE82328		
KEYWORDS	BUE82328.1	GI:23531120	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo.		
REFERENCE	1 (bases 1 to 486)		
AUTHORS	Ronald,M.F., Lennon,G. and Soars,M.B.		
TITLE	Normalization and subtraction: two approaches to facilitate gene		
JOURNAL	discovery		
PUBMED	Genome Res. 6 (9) , 791-806 (1996)		
COMMENT	8889548		
	Contact: McCray, PB		

University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel.: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@iowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open BioSystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 1-42, >AT rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

```
FEATURES
source
location/Qualifiers
1..466
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-BCL-acb-g-17-0-UI"
/tissue_type="Lung"
```

```

/def_stage="Adult and Fetal"
/def_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EC1"
/node="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGTGGCTTAC.
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_LIB=UI-CF-EC1
TAG_SEQ=AAGTGGCTTAC"

```

Query Match	2.24	Score 43	DB 5	Length 486
Best Local Similarity	47.3%	Prod. No. 0.63		
Matches 130	Conservative 0	Mismatches 145	Indels 0	Gaps 0
Qy	1389	TACACCGATCGTTGCTGCGAAA	CGTGCACAGCTTCGTGACATTTTGTGAAAA	CAAGAG 1448
Db	183	TACCCCTCATGATGATGAAA	CCAGCAACAGATCAGGCTCATTCAGAAAA	CCAGAG 242
Qy	1449	CATCACACCAAACCTTCTCTATTA	AAAGCGCACATGATTCGCGCGACGTGACATCGG	1508
Db	243	AGAGAGATATGATATTTGGAAA	CTGATATAGTCTTTTTAAAAA	CAATCTGTCAAGTTT 302
Qy	1509	CGTTGGCTACCGGACGGTAA	CCAGAGATGGGGACAGTGACATTTGGATTAAT	TCCTGAA 1568
Db	303	CATGAACAACCCCAATAA	ACMAAATGAAAAAATACATAGGATCA	CTCACCCAAACA 362
Qy	1569	CGTTGCTATGTGAAGAAGTCCAGTT	CTCATCCACACAGCAAAAAAGCAGCTACTCGTT	1628
Db	363	AATGTGCTGTGTTAACAA	CAACAGCAGTACATCATCATTTCAAA	CCATAGGTAGTTACA 422
Qy	1629	TACTGCTACTGCGCGACAGCTTTG	AAAAATCTCC	1663
Db	423	GTACAAAAAGACTGGCACCGGTGTG	ACACCCAC	457

RESULT 3				
A1269898				
LOCUS	A1269898	500 bp	mRNA	linear EST 28-JAN-1999
DEFINITION	XQ04001.x1 NCI_CGAP_kid3 Homo sapiens CDNA clone IMAGE:1867969 3'			
ACCESSION	A1269898			
VERSION	A1269898			
KEYWORDS	A1269898.1 GI:3889065			
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Homidae; Homo.			
	1 (bases 1 to 500)			
REFERENCE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .			
AUTHORS	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			
TITLE	Tumor Gene Index			
JOURNAL	Unpublished (1997)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabbs-r@mail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmerit-Buck, M.D., Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D.			

QY 1268 AACCGATGTCATGACACCGGAAAGCTTCATACATGACAGATCTTCTTG 1321
 |||||
 Db 436 AACCTGATATGTTGCACCTGCGAATATTTATACATCGGCAAGATCAATATGTTG 489
 |||||

RESULT 6
 LOCUS BJ393925
 DEFINITION BJ393925 Dictyostelium discoideum cDNA library, SP Dictyostelium
 accession BJ393925
 VERSION BJ393925
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum
 ORGANISM Dictyostelium discoideum
 Eukaryota; Eukaryota; Dictyostelid; Dictyostelium.

REFERENCE
 1 (bases 1 to 641)
 Trushikare, H., Tanaka, Y., Kohara, Y., and Shin-1, T.
 Full length cDNA of Dictyostelium discoideum at the slug stage
 JOURNAL Unpublished (2002)
 COMMENT Contact: Tadao Shin-1
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

FEATURES
 source
 1..641
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="dd83105"
 /sex="mat A"
 /dev_stage="slug stage"
 /clone_1ib="Dictyostelium discoideum cDNA library, SP"

ORIGIN
 Query Match 2.2%; Score 42; DB 3; Length 641;
 Best Local Similarity 60.5%; Pred. No. 1.4;
 Matches 69; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1208 ACAATATCAACCATGTGACAGTTCTTCACTGACGACGACCAAGATGACGATCA 1267
 |||||
 Db 169 AATAAGAAATATATGATGATCTTCAATCAAGGTCACATGATGTAATTGA 228
 |||||

QY 1268 AACCGATGTCATGACACCGGAAAGCTTCATACATGACAGATCTTCTTG 1321
 |||||
 Db 229 AACCTGATATGTTGCACCTGCGAATATTTATACATCGGCAAGATCAATATGTTG 282
 |||||

RESULT 7
 LOCUS AJ689929
 DEFINITION AJ689929 Homo sapiens cDNA clone IMAGE:2271232 3',
 mRNA sequence.
 accession AJ689929
 VERSION AJ689929
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 1 (bases 1 to 457)
 NCI-CCAP htcp://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgepds-remail.nih.gov
 Tissue Procurement: Christopher Moshak, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 www.bio.lnl.gov/btrp/image/image.html
 Insert Length: 577 Std Error: 0.00
 Seq primer: -40UP from G1bco.
 Location/Qualifiers
 1..457
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2271232"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /clone_1ib="NCI CGAP Lu24"
 /note="Organ: lung; Vector: pRT3D-Pac (pharmacia) with a
 modified polylinker; Plasmid DNA from the normalized
 library NCI-CCAP Lu5 was prepared, and 88 clones were
 made in vitro. Following HAP purification, this DNA was
 used as tracer in a subtractive hybridization reaction.
 The driver was PCR-amplified cDNAs from a pool of 5,000
 clones made from the same library (cloneids
 141920-1417991 and 1520904-1522439). Subtraction by Bento
 Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 2.2%; Score 41.8; DB 1; Length 457;
 Best Local Similarity 57.1%; Pred. No. 1.4;
 Matches 76; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 174 TGCTGCTTTCTGTGTGAATGAAAATGCAATCCCAAGGTTGCAAGAAAGCT 233
 |||||
 Db 325 TGCAAGTATGTGTGATGAAAACAGCTAATGAAACACTGATCTGATTTAAAGTTT 384
 |||||

QY 234 TGAACACGTCCCGCAATATTAATCTTCATTTATTCATTTCAATGACCAATTTTGA 293
 |||||
 Db 385 TCTACCTGTATACACTCTACTACATCTTCATCAACCAACAACTTGTGCAATTAACA 444
 |||||

QY 294 AGAAACAAACAG 306
 |||||
 Db 445 TGAGCAAAACAG 457
 |||||

RESULT 8
 LOCUS AA39583
 DEFINITION AA39583 464 bp mRNA linear EST 16-MAY-1997
 3', mRNA sequence.
 accession AA39583
 VERSION AA39583
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 1 (bases 1 to 464)
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
 Moore, B., Schellenberg, K., Sepke, M., Tan, F., Theising, B.,
 White, Y., Wyllie, T., Waterston, K., and Wilson, R.
 WashU-Merck EST Project 1997
 Unpublished (1997)
 COMMENT Contact: Wilson R.
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNLN; contact the

CDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pWT3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)₁₈ tail. The sequence tag for this library is GCTCAGAGCT.

TAG_TISSUE=chondrosarcoma
TAG_LIB=UI-H-ED0

ORIGIN

Query Match	2.2%	Score 41.8;	DB 6;	Length 628;
Best Local Similarity	57.1%	Score 41.8;	DB 6;	Length 628;

Matches	76;	Conservative	0;	Mismatches	57;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	-----	--------	----	------	----

[illegible]

234 TGAACAGTCCCGGCAATATACTCCATATTATCCAATTCATATGGACCAATTTTAGA 293

394 TCTACCTGTACACACTACTACTCCATACAAACCACTTTGGTCCAATAAACA 453

294 AGATACTATAAG 306

RESULT 12

Accession	Source	Library	Insert size	Library size	Library type	Library description
U000000000	GenBank	U000000000	878 bp	1.0 kb	Linear	U000000000
U000000001	GenBank	U000000001	878 bp	1.0 kb	Linear	U000000001
U000000002	GenBank	U000000002	878 bp	1.0 kb	Linear	U000000002
U000000003	GenBank	U000000003	878 bp	1.0 kb	Linear	U000000003
U000000004	GenBank	U000000004	878 bp	1.0 kb	Linear	U000000004
U000000005	GenBank	U000000005	878 bp	1.0 kb	Linear	U000000005
U000000006	GenBank	U000000006	878 bp	1.0 kb	Linear	U000000006
U000000007	GenBank	U000000007	878 bp	1.0 kb	Linear	U000000007
U000000008	GenBank	U000000008	878 bp	1.0 kb	Linear	U000000008
U000000009	GenBank	U000000009	878 bp	1.0 kb	Linear	U000000009
U000000010	GenBank	U000000010	878 bp	1.0 kb	Linear	U000000010
U000000011	GenBank	U000000011	878 bp	1.0 kb	Linear	U000000011
U000000012	GenBank	U000000012	878 bp	1.0 kb	Linear	U000000012
U000000013	GenBank	U000000013	878 bp	1.0 kb	Linear	U000000013
U000000014	GenBank	U000000014	878 bp	1.0 kb	Linear	U000000014
U000000015	GenBank	U000000015	878 bp	1.0 kb	Linear	U000000015
U000000016	GenBank	U000000016	878 bp	1.0 kb	Linear	U000000016
U000000017	GenBank	U000000017	878 bp	1.0 kb	Linear	U000000017
U000000018	GenBank	U000000018	878 bp	1.0 kb	Linear	U000000018
U000000019	GenBank	U000000019	878 bp	1.0 kb	Linear	U000000019
U000000020	GenBank	U000000020	878 bp	1.0 kb	Linear	U000000020
U000000021	GenBank	U000000021	878 bp	1.0 kb	Linear	U000000021
U000000022	GenBank	U000000022	878 bp	1.0 kb	Linear	U000000022
U000000023	GenBank	U000000023	878 bp	1.0 kb	Linear	U000000023
U000000024	GenBank	U000000024	878 bp	1.0 kb	Linear	U000000024
U000000025	GenBank	U000000025	878 bp	1.0 kb	Linear	U000000025
U000000026	GenBank	U000000026	878 bp	1.0 kb	Linear	U000000026
U000000027	GenBank	U000000027	878 bp	1.0 kb	Linear	U000000027
U000000028	GenBank	U000000028	878 bp	1.0 kb	Linear	U000000028
U000000029	GenBank	U000000029	878 bp	1.0 kb	Linear	U000000029
U000000030	GenBank	U000000030	878 bp	1.0 kb	Linear	U000000030
U000000031	GenBank	U000000031	878 bp	1.0 kb	Linear	U000000031
U000000032	GenBank	U000000032	878 bp	1.0 kb	Linear	U000000032
U000000033	GenBank	U000000033	878 bp	1.0 kb	Linear	U000000033
U000000034	GenBank	U000000034	878 bp	1.0 kb	Linear	U000000034
U000000035	GenBank	U000000035	878 bp	1.0 kb	Linear	U000000035
U000000036	GenBank	U000000036	878 bp	1.0 kb	Linear	U000000036
U000000037	GenBank	U000000037	878 bp	1.0 kb	Linear	U000000037
U000000038	GenBank	U000000038	878 bp	1.0 kb	Linear	U000000038
U000000039	GenBank	U000000039	878 bp	1.0 kb	Linear	U000000039
U000000040	GenBank	U000000040	878 bp	1.0 kb	Linear	U000000040
U000000041	GenBank	U000000041	878 bp	1.0 kb	Linear	U000000041
U000000042	GenBank	U000000042	878 bp	1.0 kb	Linear	U000000042
U000000043	GenBank	U000000043	878 bp	1.0 kb	Linear	U000000043
U000000044	GenBank	U000000044	878 bp	1.0 kb		

CM901981.1 GT:56444251

WORDS	GSS.
SOURCE	BOB taurus (cow

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Theria; Laurasiatheria; Cetartiodactyla. Ruminantia.

REFERENCE 1 (bases 1 to 678)
Larkin D M Denton V Rothman B Ryan V Yliniemi J and

TITLE End sequencing of Holstein BAC library RPCI-42

Contact: Harry's Lewin

University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA

Fax: 217 244 5617
Email: b-lewin@uiuc.edu

Source

availability, please contact

availability, please contact

(pdejong@mail.cho.org). Clone
Resources (<http://www.choi>

Funding was provided in part by the National Institutes of Health (NIH) (R01AG023401) and the Department of Defense (W81XWH-04-2-0002). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or the Department of Defense.

USDA-CSREES (Livestock Genom
AG58-5438-2-313 from USDA-AF

Plate: 139 row: 5 column:

Seq primer: SP6
Class: BAC ends.

COURSES	LOCATIONS	QUALIFIERS

```

SOURCE
L. : 6/8
/organism="Bos taur

```

```
/mol_type="genomic"
/attr{sn="Holoferon"
```

```

/occurance: none
/db_xref="taxon:991

```

/clone="RPC142_139B1."
/sex="male"
/cell_type="blood"
/clone_id="RPC1-42"
/note="Vector: DBACE3.6; Site 1: BcORI; Site 2: BcORI;
RPC1-42 Bovine BAC library (Male) produced by Pieter de
Jong"

ORIGIN

Query Match 2.2%; Score 41.8; DB 10; Length 678;
Best Local Similarity 56.0%; Pred. No. 1.6;
Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 197 AAAATGGAACCTCCAAAGTTTGGAGAAAGCTTGAACAGTCCCGCAATATA 256
DB 491 AAAAGAAAACTATTAAAGAAATGAGACATTCAGAAACCTCCAGACAAATTTA 432
QY 257 AACTCATATTATTCATTCATATGAGACCAATTTTAGAAGAAACAGCAGCTGGAA 316
DB 431 AAGCCCAACGTTCAATCATAGAGTCCCAAGAGAGACAGAAAGACCATGAGA 372
QY 317 AAACAGGGGCAAGATTCTCG 337
DB 371 AATACTTGAGAGATATATG 351

RESULT 13
CWS36661/c 685 bp DNA linear GSS 08-OCT-2004
LOCUS CWS36661/c
DEFINITION OP_Ba0033K11.r OP_Ba Oryza punctata genomic clone OP_Ba0033K11
3' genomic survey sequence.

ACCESSION CWS36661
VERSION CWS36661.1 GI:54010883
KEYWORDS GSS.
SOURCE Oryza punctata
ORGANISM Oryza punctata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 685)
Samiguel, P., Westerman, R., Kim, H., Yu, Y., Wisotski, M., Yost, D.,
Stum, D., Rao, K., Luo, M., Jettly, R., Kudrna, D., Miller, C.,
Hatfield, J., Soderlund, C., Wang, R. and Jackson, S.A.
Hatfield, J., Soderlund, C., Wang, R. and Jackson, S.A.
OMP Project - Purdue University
Unpublished (2004)
Contact: Scott A. Jackson
Jackson Laboratory
Purdue University
915 W. State St., West Lafayette, IN 47907, USA
Tel: 7854963621
Fax: 7854967255
Email: sjackson@purdue.edu

TITLE
JOURNAL
COMMENT

Basecalling by phred version 0.020425.c. This sequence was derived
from the raw sequence read by clipping with Lucy version 1.19e.
Bases 104-788 of the raw sequence (length 1413) were retained after
clipping.
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Insert Length: 161000 Std Error: 0.00
Place: 0033 row: K column: 11
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1..685
/organism="Oryza punctata"
/mol_type="genomic DNA"
/db_xref="taxon:4537"
/clone="OP_Ba0033K11"
/issue_type="young leaves"
/lab_host="DH10B-T1 phage resistant"
/clone_id="OP_Ba"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 2.2%; Score 41.8; DB 10; Length 685;
Best Local Similarity 49.3%; Pred. No. 1.6;
Matches 109; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 156 GCAGGGGACAGCTGCGTCTGCTCTTTCTGCGGAATCGAAATGGAATCTCCCAA 215
DB 615 GCAGATGTTAAAGGGGCTCTTGAATTAATCTTCAACAGCAATGTCAGAGTGGCAT 556
QY 216 AGTTTGCAGAGAGACCTTGAACAGTCCCGCAATTAATCACTCATTTTCAAT 275
DB 555 CAGCATTTAGAAACAGGCGCTTCAACTCAAGCACTTGCAGAAAGTCACTACCTGT 496
QY 276 CAATGACCAATTTAGAAAGAAACAACAGCAGCTGGAAGAAACAGGGGCAAGATTCT 335
DB 495 ACAATTTCAACCTTTGTAAGAAAGAAATCAGATTTAGAGAAATACATTGCAACATA 436
QY 336 CCACTACATACCTGATTAATGCTTACATTTGTCAGATAGAG 376
DB 435 CTTATATCAACCAATCATGCAAGATTAATCAATATGTAAG 395

RESULT 14
HSM805591/c 7301 bp mRNA linear HTC 22-SEP-2004
LOCUS HSM805591/c
DEFINITION Homo sapiens mRNA; cDNA DKFZp434J1323 (from clone DKFZp434J1323).
ACCESSION AL834204
VERSION AL834204.1 GI:21739712
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE

1 (bases 1 to 7301)
Blum, H., Bauerbachs, S., Mewes, H.W., Weil, B., Amid, C., Oeanger, A.,
Fobo, G., Han, W. and Wiemann, S.
The German cDNA Consortium
Direct Submission
Submitted (12-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by LMU (Ludwig Maximilians University, Munich/Germany)
within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp434J1323) is available at the RZPD Deutsche
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp434J1323
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

COMMENT

FEATURES
source

Location/Qualifiers
1..7301
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp434J1323"
/db_xref="taxon:9606"
/clone="DKFZp434J1323"
/issue_type="testis"
/clone_id="434 (synonym: hce83). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="ankyrin repeat-containing protein 12"
1..7301
/gene="DKFZp434J1323"
/join(193..3069,3069..6311)
/gene="DKFZp434J1323"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH56382.1"
/db_xref="GI:52545871"
/translation="MPKSGFTYPIQSENSDSDSNVKEPYGRKSKDKIASYSKTPKIE"

gene

CDS

QY 181 TTTCTGTGGAATCTGAAAAATGAAACTCCCAAAAGTTTGAGAGAAAGCTTGAACA 240
DB 181 TTTCTGTGGAATCTGAAAAATGAAACTCCCAAAAGTTTGAGAGAAAGCTTGAACA 240
QY 241 GTCCCGCAATATATATACTCCATATTATCCAAATTCATGAAACCAATTTTAAAGAAACA 300
DB 241 GTCCCGCAATATATATACTCCATATTATCCAAATTCATGAAACCAATTTTAAAGAAACA 300
QY 301 AAAACGACGTGAAAAAACAAGGGGCAAGATTCTGACCTACCTACCTGATTTATGCTTAC 360
DB 301 AAAACGACGTGAAAAAACAAGGGGCAAGATTCTGACCTACCTACCTGATTTATGCTTAC 360
QY 361 ATTGTGAGATATAGAGGCGATGTTAAGTACGAAACAAGACCAATGAGACGCTGAATCC 420
DB 361 ATTGTGAGATATAGAGGCGATGTTAAGTACGAAACAAGACCAATGAGACGCTGAATCC 420
QY 421 GTGAGACCTTATTTTCCGATATACAGAAATAGATCCCAAGCTTTTCAAAAAGGGGATCA 480
DB 421 GTGAGACCTTATTTTCCGATATACAGAAATAGATCCCAAGCTTTTCAAAAAGGGGATCA 480
QY 481 GAGCTTGTAAAGCAGTGGGCGTGTATACAAAGCAGAAAAATTAAGAGTGCATTTAAGA 540
DB 481 GAGCTTGTAAAGCAGTGGGCGTGTATACAAAGCAGAAAAATTAAGAGTGCATTTAAGA 540
QY 541 GGCATCGAACCAATCGCAAAATTCGCAATAGCAATGATGTCTATATATTAACGGCAAG 600
DB 541 GGCATCGAACCAATCGCAAAATTCGCAATAGCAATGATGTCTATATATTAACGGCAAG 600
QY 601 CCTGAGATATAGATGATGAATGTTGGCGCTGGAATTTGCAAAAGCGATGTGCTCAG 660
DB 601 CCTGAGATATAGATGATGAATGTTGGCGCTGGAATTTGCAAAAGCGATGTGCTCAG 660
QY 661 AGCAGCTACGGGTTGTATGACAAAGACAGATCGTAGGGGTTGCCGATACAGGGCTGAT 720
DB 661 AGCAGCTACGGGTTGTATGACAAAGACAGATCGTAGGGGTTGCCGATACAGGGCTGAT 720
QY 721 ACAGGTCGCAATGACAGTTGATGATGAGAGCTTCCGCGGGAATTAATCTGCAATTAAT 780
DB 721 ACAGGTCGCAATGACAGTTGATGATGAGAGCTTCCGCGGGAATTAATCTGCAATTAAT 780
QY 781 GCAATGGGACCGACCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCAATGGGACCGACCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGATATTAGAAACGGCTCCACTAATAAGGAATGGCGCTCAGGCGAATCTAGCTTTC 900
DB 841 TCCGATATTAGAAACGGCTCCACTAATAAGGAATGGCGCTCAGGCGAATCTAGCTTTC 900
QY 901 CAATCTATCATGATAGCGGTGGGGAAGCTTGGAGACTTACCTTCGATCTGCAACCTTA 960
DB 901 CAATCTATCATGATAGCGGTGGGGAAGCTTGGAGACTTACCTTCGATCTGCAACCTTA 960
QY 961 TTCAGCCCAAGCATACAGTGTGTGTGTCGCAAACTCTGCGGAGACGACAGTGT 1020
DB 961 TTCAGCCCAAGCATACAGTGTGTGTGTCGCAAACTCTGCGGAGACGACAGTGT 1020
QY 1021 AATGGGCGCTTACACACAGATTCGAGAAATGTGAATGCTATGTGCGCAAAAATGATATG 1080
DB 1021 AATGGGCGCTTACACACAGATTCGAGAAATGTGAATGCTATGTGCGCAAAAATGATATG 1080
QY 1081 ACGATCCTTTTCTGCTGCGGGAATGAGACCGAACGCGGAACCATGACGACCAAGGC 1140
DB 1081 ACGATCCTTTTCTGCTGCGGGAATGAGACCGAACGCGGAACCATGACGACCAAGGC 1140
QY 1141 ACAGCTTAAATATGCAATACAGTGTGAGCTACGAAAACTTCCGCGCAAGCTTTGGGCTT 1200
DB 1141 ACAGCTTAAATATGCAATACAGTGTGAGCTACGAAAACTTCCGCGCAAGCTTTGGGCTT 1200
QY 1201 TATGCGGACCAATATCAACATGTGTGCAAGTTCTTCTTCACTGTGACCAAAAGGATGA 1260
DB 1201 TATGCGGACCAATATCAACATGTGTGCAAGTTCTTCTTCACTGTGACCAAAAGGATGA 1260

QY 1261 CGGATCAAAACCGGATGTATGCAACCGGGAAAGTTTCACTATGACGAAGAATCTTCTCTT 1320
DB 1261 CGGATCAAAACCGGATGTATGCAACCGGGAAAGTTTCACTATGACGAAGAATCTTCTCTT 1320
QY 1321 GCAACCGGATTCCTCTCTTGGGCGGAACCATGACAGTAATATGCAATCATGGGTGAACG 1380
DB 1321 GCAACCGGATTCCTCTCTTGGGCGGAACCATGACAGTAATATGCAATCATGGGTGAACG 1380
QY 1381 TCCATGGCTACACCGATGTGTGTGTGAAAAAGTGGCACGCTTCTGTGACATTTTGTGAAA 1440
DB 1381 TCCATGGCTACACCGATGTGTGTGTGAAAAAGTGGCACGCTTCTGTGACATTTTGTGAAA 1440
QY 1441 AACAGAGGCTACACCAAAAGCTTCTCTATTAATAAGGGGACATGATGGCGGTCAGCT 1500
DB 1441 AACAGAGGCTACACCAAAAGCTTCTCTATTAATAAGGGGACATGATGGCGGTCAGCT 1500
QY 1501 GACATCGGCTTTGGCTACCCGAAACGTTAACCAAGATGGGACGAGTGAATTAATA 1560
DB 1501 GACATCGGCTTTGGCTACCCGAAACGTTAACCAAGATGGGACGAGTGAATTAATA 1560
QY 1561 TCCCTGAACGTTGCTATGTGTAACGAGTCCAGTTCTATTCACACGCAAAAAGCGACG 1620
DB 1561 TCCCTGAACGTTGCTATGTGTAACGAGTCCAGTTCTATTCACACGCAAAAAGCGACG 1620
QY 1621 TACTGTTTACTGTACTGCGCGCAAGCTTTGAAAAATCTCCCTGATGATGGTGTGATGCC 1680
DB 1621 TACTGTTTACTGTACTGCGCGCAAGCTTTGAAAAATCTCCCTGATGATGGTGTGATGCC 1680
QY 1681 CTGCGAGCAACACTGCTTCCGTAAGCTTGTCAATGATCTGACCTTGTCAATTAACGCT 1740
DB 1681 CTGCGAGCAACACTGCTTCCGTAAGCTTGTCAATGATCTGACCTTGTCAATTAACGCT 1740
QY 1741 CCAATATGGCACACGATATGTAAGAAATGAATGATGATGATGATGATGATGATGATGAT 1800
DB 1741 CCAATATGGCACACGATATGTAAGAAATGAATGATGATGATGATGATGATGATGATGAT 1800
QY 1801 GCGCGCATATAACGTAGAAATATGATTTATTAATGACCAACAAAGCGGGAAGTATCAATT 1860
DB 1801 GCGCGCATATAACGTAGAAATATGATTTATTAATGACCAACAAAGCGGGAAGTATCAATT 1860
QY 1861 GAGGTACAGGCTTATTAACGTACCGGTTGGAACCAAGACCTTCTGTGGCAATTTGTAAT 1920
DB 1861 GAGGTACAGGCTTATTAACGTACCGGTTGGAACCAAGACCTTCTGTGGCAATTTGTAAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923

RESULT 2
US-09-920-954-5
Sequence 5, Application US/09920954
Patent No. 6739228
GENERAL INFORMATION:
APPLICANT: TAKAIIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKELI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUW
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT FILING DATE: US/09/920, 954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509, 814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24

QY 1261 CGGATCAACCGGATGTCATGCGACCGGGAACCTTGATATCTACAGAAAGATCTTCTCTT 1320
DB 1261 CGGATCAACCGGATGTCATGCGACCGGGAACCTTGATATCTACAGAAAGATCTTCTCTT 1320
QY 1321 GCAACGATTCCTCTCTTCTGCGGCAACATGACAGTAATATGATACATGCGGTGAGCG 1380
DB 1321 GCAACGATTCCTCTCTTCTGCGGCAACATGACAGTAATATGATACATGCGGTGAGCG 1380
QY 1381 TCCATGCGTCAACCGATTCGTTGCTGGAACGTTGCGACAGCTTCGTGAGCAATTTTGAAA 1440
DB 1381 TCCATGCGTCAACCGATTCGTTGCTGGAACGTTGCGACAGCTTCGTGAGCAATTTTGAAA 1440
QY 1441 AACAGAGGATCAACCAAGCCTTCTATTAAAGGCGGCACTGATTCGCGGTGAGCT 1500
DB 1441 AACAGAGGATCAACCAAGCCTTCTATTAAAGGCGGCACTGATTCGCGGTGAGCT 1500
QY 1501 GACATCGGCTTGCTGCTACCGCAACGCTTAACAGATGCGGAGAGTGAATTGATTA 1560
DB 1501 GACATCGGCTTGCTGCTACCGCAACGCTTAACAGATGCGGAGAGTGAATTGATTA 1560
QY 1561 TCCCTGAAAGTTGCTGCTATGTAACAGTCAAGTCTTATCCAACGCAAAAGCGAGC 1620
DB 1561 TCCCTGAAAGTTGCTGCTATGTAACAGTCAAGTCTTATCCAACGCAAAAGCGAGC 1620
QY 1621 TACTGCTTACTGCTACTGCGCGCAAGCCTTGAAATCTCCCTGCTATGCTGATGCC 1680
DB 1621 TACTGCTTACTGCTACTGCGCGCAAGCCTTGAAATCTCCCTGCTATGCTGATGCC 1680
QY 1681 CCTGCGAGCAACCTGCTTCGTTACGTTGCAATGATCTGAGCCTTGTCAATACCGT 1740
DB 1681 CCTGCGAGCAACCTGCTTCGTTACGTTGCAATGATCTGAGCCTTGTCAATACCGT 1740
QY 1741 CCAATGCGACACATGATGTAAGAAATGCTTACTGCGCAATGATTAACCTGAGT 1800
DB 1741 CCAATGCGACACATGATGTAAGAAATGCTTACTGCGCAATGATTAACCTGAGT 1800
QY 1801 GCGCGCAATTAACGTAAGAAATGATTTATTAATGACACCAAGGCGGAGCTATCAAT 1860
DB 1801 GCGCGCAATTAACGTAAGAAATGATTTATTAATGACACCAAGGCGGAGCTATCAAT 1860
QY 1861 GAGGTACAGGCTTATTAACGTAACGCTTGACCAACAGACCTTCTCGTTGCAATTTGAT 1920
DB 1861 GAGGTACAGGCTTATTAACGTAACGCTTGACCAACAGACCTTCTCGTTGCAATTTGAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923

RESULT 5
US-09-509-814A-3

Sequence 3, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKETI, KATSUNISHI
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEIYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509, 814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin version 3.0
SEQ ID NO 3

QY 10 AAGAAAAAGGTTTATCTGTTTATCTGATGACGAGGATTTGTCGATCTGTTGCGTTA 69
DB 7 AAGAAAGAGGTTTATCTGTTTATCTGATGACGAGGATTTGTCGATCTGTTGCGATTA 66
QY 70 AGTATCCATCGGAGGAGGTCGAAAGAAATTTGATCGATGATTTTCAAGAAATTCAGCA 129
DB 67 AACATCCCTCGGCTGATGATGCAAGGATTTTATCTGATGATTTTAAAGAAATTCAGCA 126
QY 130 AACATGATGCTAAAGTTTCTCAAGCAAGGAGGAGGATGCTGCTGCTTTCTGAGT 189
DB 127 AACATGATGCTAAAGTTTCTCAAGCAAGGAGGAGGATGCTGCTGCTTTCTGAGT 186
QY 190 GATTTGAAATGTAATCTCCAAAGGTTTCAAGAAAGTTTGAAGAAAGTTGAAACAGTCCGCA 249
DB 187 GAGTCTGAAATGTAATCTTAAAGGATTTGCTAAAGAAATTTGAAACAGTCCGCA 246
QY 250 AATATTAATCTCAATTTATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 309
DB 247 AATATTAATCTCAATTTATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 306
QY 310 CTGAAAAAAGGAGGAGGAGGATTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 369
DB 307 CTGAAAAAAGGAGGAGGAGGATTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 366
QY 370 TATAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 429
DB 367 TATAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 426
QY 430 TATTTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489
DB 427 TATTTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 486
QY 490 AAGCAAGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 549
DB 487 AAGCAAGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
QY 550 CAATGCGCAATTCGATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 609
DB 547 CAATGCGCAATTCGATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 606
QY 610 AAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
DB 607 AAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
QY 670 GGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
DB 667 GGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
QY 730 AATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
DB 727 AATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
QY 790 CGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
DB 787 CGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
QY 850 GGAAGGCTGCTCAATTAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 909
DB 847 GGAAGGCTGCTCAATTAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
QY 910 ATGATAGCGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 969

Db 907 ATGGATAGCAGTGTGGCTGGAGGCTTCCATTCGCAAACTTATTCACGCA 966
Qy 970 GCATACAGTGTGTGGCTGGAGGCTTCCATTCGCAAACTTATTCACGCA 1029
Db 967 GCATACAGTGTGTGGCTGGAGGCTTCCATTCGCAAACTTATTCACGCA 1026
Qy 1030 TACACAAGATTCAGAAATGTGATGATATGTGCGCAAAATGATGACATCTT 1089
Db 1027 TACACAAGATTCAGAAATGTGATGATATGTGCGCAAAATGATGACATCTT 1086
Qy 1090 TTGCGTGTGGGGAATGAGACCGGACCGGACCAATCAGTGCACGACGACGCTAA 1149
Db 1087 TTGCGTGTGGGGAATGAGACCGGACCGGACCAATCAGTGCACGACGACGCTAA 1146
Qy 1150 AATGGAATACAGTGTGGAGTACGAAACCTCGGCAAGCTTGGTCTTATGGGAC 1209
Db 1147 AATGGAATACAGTGTGGAGTACGAAACCTCGGCAAGCTTGGTCTTATGGGAC 1206
Qy 1210 AATATCAACATGTGCAAGTCTCTTCACTGCAACCGCAAGATGACGATCAAA 1269
Db 1207 AATATCAACATGTGCAAGTCTCTTCACTGCAACCGCAAGATGACGATCAAA 1266
Qy 1270 CCGGATGTCAATGCAACCGGGAACGTTCAATATCAAGAAATCTTCTTGCACCGAT 1329
Db 1267 CCGGATGTCAATGCAACCGGGAACGTTCAATATCAAGAAATCTTCTTGCACCGAT 1326
Qy 1330 TCTCTCTTCTGGGCGCAACATGACATTAATATGATATGATGATGATGATGATGAT 1389
Db 1327 TCTCTCTTCTGGGCGCAACATGACATTAATATGATGATGATGATGATGATGATGAT 1386
Qy 1390 ACAACGATGTTGTGGAACAGTGCAGCTCTGAGCACTTTTGTGAAAACAGAGGC 1449
Db 1387 ACAACGATGTTGTGGAACAGTGCAGCTCTGAGCACTTTTGTGAAAACAGAGGC 1446
Qy 1450 ATCAACCAAGCTTCTCTATTAAGGCGCACTGATGCGGTCAGCTGACATCGGC 1509
Db 1447 ATCAACCAAGCTTCTCTATTAAGGCGCACTGATGCGGTCAGCTGACATCGGC 1506
Qy 1510 CTGCTGATCCGCAACGTTAATCAAGATGCGGTCAGCTGATGCGGTCAGCTGAT 1569
Db 1507 CTGCTGATCCGCAACGTTAATCAAGATGCGGTCAGCTGATGCGGTCAGCTGAT 1566
Qy 1570 GTTGCCTATGTGAACGATCTCTATTAAGGCGCACTGATGCGGTCAGCTGAT 1629
Db 1567 GTTGCCTATGTGAACGATCTCTATTAAGGCGCACTGATGCGGTCAGCTGAT 1626
Qy 1630 ACTGCTACTGCGGCAAGCTTGTGAATCTCTGATGATGATGATGATGATGATGAT 1689
Db 1627 ACTGCTACTGCGGCAAGCTTGTGAATCTCTGATGATGATGATGATGATGATGAT 1686
Qy 1690 ACAACGCTCTGCAACGTTGATGATGATGATGATGATGATGATGATGATGAT 1749
Db 1687 ACAACGCTCTGCAACGTTGATGATGATGATGATGATGATGATGATGATGAT 1746
Qy 1750 ACAACGCTCTGCAACGTTGATGATGATGATGATGATGATGATGATGATGAT 1809
Db 1747 ACAACGCTCTGCAACGTTGATGATGATGATGATGATGATGATGATGATGAT 1806
Qy 1810 AACGTAAGAAATGATATTAATGATGATGATGATGATGATGATGATGATGATGAT 1869
Db 1807 AACGTAAGAAATGATATTAATGATGATGATGATGATGATGATGATGATGATGAT 1866
Qy 1870 GCTTATACGTAACGCTGTGACCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1923
Db 1867 GCTTATACGTAACGCTGTGACCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1920

RESULT 6
US-09-920-954-3
; Sequence 3, Application US/0920954
; Patent No. 6759228
; GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/J98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 3
LENGTH: 1920
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
US-09-920-954-3
Query Match 76.8%; Score 1477.2; DB 3; Length 1920;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
Qy 10 AAAAAAGGTTGTTTATCTGTTTATCACTGACGCAATTTGTGCACTGTGCTTA 69
Db 7 AAAAAAGGTTGTTTATCTGTTTATCACTGACGCAATTTGTGCACTGTGCTTA 66
Qy 70 AATAATCCATCGCAGTGTGCAAGAAATTTGATCTGGAATTTCAAGAAATTCAGACA 129
Db 67 AATAATCCATCGCAGTGTGCAAGAAATTTGATCTGGAATTTCAAGAAATTCAGACA 126
Qy 130 ACAACGATGTTGTAAGTTTCTCAAGCAGGCGCAGATGCTGCTGCTGCTGCTGCTG 189
Db 127 ACAACGATGTTGTAAGTTTCTCAAGCAGGCGCAGATGCTGCTGCTGCTGCTGCTG 186
Qy 190 GAATCTGAATATGTGAACCTCCCAAGGTTTGAGAAAGACCTTGAAACATCTCCGCA 249
Db 187 GAATCTGAATATGTGAACCTCCCAAGGTTTGAGAAAGACCTTGAAACATCTCCGCA 246
Qy 250 AATAATTAATCTCATATTAATCAATTCATGACCAATTTTGAAGAAACAAACAGCAG 309
Db 247 AATAATTAATCTCATATTAATCAATTCATGACCAATTTTGAAGAAACAAACAGCAG 306
Qy 310 CTGAAAAAAGCAGGCGCAAGATTTCTGACATACCTGATTTATGCTTAATGCTGAG 369
Db 307 CTGAAAAAAGCAGGCGCAAGATTTCTGACATACCTGATTTATGCTTAATGCTGAG 366
Qy 370 TATGAGGCGATGTTAAGTCAAGCAAGCAATTCATGACAGCTGGAATCCGTGAGCCT 429
Db 367 TATGAGGCGATGTTAAGTCAAGCAAGCAATTCATGACAGCTGGAATCCGTGAGCCT 426
Qy 430 TATTTGCGGATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 489
Db 427 TATTTGCGGATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 486
Qy 490 AAAAGAGTGGCGCTGATACAAAGCAAGAAATTAAGAGTGCATTAAGAGGCAATCGAA 549
Db 487 AAAAGAGTGGCGCTGATACAAAGCAAGAAATTAAGAGTGCATTAAGAGGCAATCGAA 546
Qy 550 CAATTCGCAATTCGCAATTAAGCAATGATGCTATTAATTAAGCAAGGCTGAGTAT 609
Db 547 CAATTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606

QY	610	AAGGTGATGTAAGAATGTTGGCGGTGGAATTGTCAAAGCGGAATGTGGCTCAGAGCAGCTAC	669
DB	607	AAGGTGATGTAAGAATGTGGCCAGAGGGATATTGTCAAAGCGGAATGTGGCCAGAGCAGCTAC	666
QY	670	GGGTTGTATGGAACAAGGACAGATCGTACGGGTTGCCGATACAGAGGCTTGATACAGGTGCG	729
DB	667	GGTTTGTATGGAACAGGCGCAGATTTGTCCGAGTTGCCGATACCTGGATTGGATACAGGAGA	726
QY	730	AATGACAGTTTCGATGATGAGCCCTTCCGCGGGGAAAATTATCTGCATTATATGCAATTGGGA	789
DB	727	AACGACAGTTTCGATGATGAGCCCTTCCGCGGGGAAAATTATGCACTATATGCACTGGGT	786
QY	790	CGGACGGAATTAATGCCAATATACGAATGGTCTATGGTAGCAGATGTGGCTGGCTCCGATTA	849
DB	787	CGGACGGAATTAATGCCAATATACGAATGGTCTATGGTAGCAGATGTGGCTGGCTCCGATTA	846
QY	850	GGAAAACGGCTCCACTAATTAAGGAATAGGCGCTCAGGCGCAATCTAGTCTTCCAACTATAC	909
DB	847	GGAAAATGGCGCAACGAATTAAGGAATGGCACTCAGCGCAATCTGGTTTTTTCAAATCATC	906
QY	910	ATGATATAGCGGTGGGGGACCTTGAGAGGACTTACCTTCGAATCTGCAAACTTATTTACGCCAA	969
DB	907	ATGATATAGCAAGTGTGGGCTTGAGAGGCTTGCTTCGAATCTGCAAACTTATTTACGCCAA	966
QY	970	GCATACAGTGTCTGTGTCCGAAATTCATACAACTCTCGGGGAGCAGCAGTGAATGGGCT	1022
DB	967	GCATTCAGTGTGAGGTGCCGAAATTCATACAACTCTCGGGGGGAGCAGCAGTGAATGGGGCC	1020
QY	1030	TACACAAACGATTTCCGAAATGTGATGACTATGTGGCGAAATAATGATATGACGATCTT	1089
DB	1027	TACACGACGATTTCCGAAATGTGATGACTATGTGAGGAAATAATGATATGACGATCTT	1088
QY	1090	TTTCGCTGCGGGGAATGAAAGAACCGAAACGCGCGAACCATCAGTGCACAGCGCACAGCTTAA	1144
DB	1087	TTTCGGGGCTGGGAATGAAGAAGCGCAACGCGGGATACATCAGTGCACCTGTACGGTAA	1144
QY	1150	AATGCAATACAGTGTGAGACTACGAAAACCTTCGCGCCAGCTTTGGGTCTTATGCGCAC	1200
DB	1147	AACGCAATACAGTGTGAGCGCGCAACCGAAAACCTGGCGTCCAAAGCTTCGATCTATGCAAGT	1206
QY	1210	AATATCAACAGTGTGCAAGTTCCTTCTTCAAGTGAACGACAAAGATGGAAGGATCAAA	1266
DB	1207	AATATTTAACAGTGTGCAAGTTCCTTCCGCGGCGCGACAAAGATGGGCGAATCAAG	1266
QY	1270	CCGATATGTCATGGCAACCGGGAACGTTCACTATCAGACGAATCTTCTTGCACCGGAT	1322
DB	1267	CCTGATGTCAATGGCGCGCAGGACATACATTTTATCAGCAAGATCTTCTTGCACCGGAT	1322
QY	1330	TCCTCTCTTCTGGGCGCAACATGACAGTAAATATGCAATACATGAGTGGAAAGCTTCATGGCT	1388
DB	1327	TCCTCTCTTCTGGGCGCAATCATGACAGCAAAATATGCTTATATGGTGGAAAGCTTCATGGCA	1388
QY	1390	AACCGATTCGTTGCTGGAAACGTGGCACAGCTTCGAGACATTTTGTGAAAACAGAGCC	1444
DB	1387	AACCGATTTGTGGCGGAATGTGCACAGCTCCGAGACATTTTGTGAAAATATGAGGA	1444
QY	1450	ATCACACCAAAAGCTTCTCTATTTAAAGCGGACAGTATGGCGGTGCAAGCTGACATCGGC	1505
DB	1447	ATCACCTCCAAAGCTTCTCTATTTGAAAGCAGCTTTGATGGCAGTGTCTGTATGTGGA	1500
QY	1510	CTTGGCTACCGCAACGGTAAACCAAGATGGGGAACGATGACATTTGATTAATTCCTGAAC	1566
DB	1507	TTGGGTTATCCGAACGGAACCAAGGATGGGGCCGAGTGAACCTGGATTAATTCGTTGAAC	1566
QY	1570	GTTGGCTATGTGAACGATTCAGTTCTCTATTCACCAAGCAAAAACGACGATATCGTTT	1622
DB	1567	GTTGGCTATGTGAACGAATTCAGTGTCCCTATACCAATGCAAAAACGACGATATATCCTT	1620
QY	1630	ACTGCTATCGCGGCAAGCTTTGAAAATCTCCCTGTATATGGCTGATATGCCCTGACAGC	1688
DB	1627	ACTGCAACGGCGGGCAAGCATTGAAAATCTCCCTGTATATGGTGGATATGCCCTGACAGC	1688
QY	1690	ACAACGTCTTCCGTAAACGGCTGTCAATGATCTGGAACCTTGTCAATTAACGCTCCAAATGGC	1741

```

Db      1667  ACAGAGTATGTAGGAAATGACTTTACTTGGCCATACATGATTAACCTGGATGCGCAAT 17466
QY      1750  ACACAGTATGTAGGAAATGACTTTACTTGGCCATACATGATTAACCTGGATGCGCAAT 18005
Db      1747  ACAAGATATGTGCGGAATGACTTTCGACGACCATTTGACATPAACTGGGATGCGCGCAAT 18006
QY      1810  AACGTAGAAATGTATTATTATTAATGACACCAAAAGCGGACGTATACATTTGAGGTACAG 18666
Db      1807  AACGTAGAAATGTATTATTATTAATGATTCGCCCAAAAGTGGAAATATATACATTGAGGTGCA 18666
QY      1870  GCTTATACGTACGCGGTGGACCAAGACCTTTCGTTGGCAATGTGAATPAA 1923
Db      1867  GCATATATATGTGCGGTGGACCAACAAACTTCGTGTTGGCAATGTGAATPAA 1920

RESULT 7
US-08-873-479-41
/ Sequence 41, Application US/08873479
/ Patent No. 5891701
/ GENERAL INFORMATION:
/ APPLICANT: Siona, Alan
/ APPLICANT: Lynne, Christlanson
/ TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
/ TITLE OF INVENTION: Having Protease Activity
/ NUMBER OF SEQUENCES: 57
/ CORRESPONDENCE ADDRESSES:
/ ADDRESS: No. 5891701o No. 5891701disk of No. 5891701th America
/ STREET: 405 Lexington Avenue
/ CITY: New York
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10174
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/873.479
/ FILING DATE: 12-JUN-1997
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Agila, Cheryl H
/ REGISTRATION NUMBER: 34,086
/ REFERENCE/DOCKET NUMBER: 5251,000-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-867-0123
/ TELEFAX: 212-878-9655
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 41:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3003 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: Genomic DNA
/ US-08-873-479-41

Query Match      60.7%; Score 1166.4; DB 2; Length 3003;
Best Local Similarity 76.0%; Pred. No. 0;
Matches 1453; Conservative 0; Mismatches 456; Indels 3; Gaps 1.

QY      12  GAAAAAGGTGTTTATCTGTTTATATGAGCGAGCGAATTTGTCGACTGTGGCTTAAG 71
Db      863  GAAAGGGGTTTTTATTCGTTTATTCAGTGTGCTGCACTATATGCTTCTGTTGCTTTAAG 922
QY      72  TAATTCATCTGCAGGTGTGCAAGGAATTTGATCTGATTTCAAGGAATTCAGACAAAC 131
Db      923  CAGTCTTCTACTATTTGGGGGGAACAATTTTGAATGGACTTTTAAAGGGGATAGAGACT 982
QY      132  AACTGATCTTAAAGGTTTCTCCAAAGCAGGGGCAACTGTGCTGCTGCTTTCTTGTGTCGA 191

```

Db 983 TACGATAGAGAGGCTGCCAACAAGAGAAAAACGGAAAGGCACTTTTCTTGTA 1042
Qy 192 ATCGAAATGTGAACTCCCAAAAGGTTGAGAAAGCTTGAACACTCCCGGAAA 251
Db 1043 CTCTGAAATGTGAAATCCCAAGAGTATTCAAAACAACTAGATGTCCAGGGA 1102
Qy 252 TAAATACTCATATTATCCAAATTCATATGACCAATTTAGAGAAAACAAACAGCAGCT 311
Db 1103 TACAGCTATATATGCTTCAATTTGAGGACCTATTTTAGAGAAAACGAACTTCACT 1162
Qy 312 GAAAAAACAAGGGCAAAAGTTCTGCACTATCTGATTTATGCTTCAATTTGAGTA 371
Db 1163 AGGAAGAAGCGGAGCGAAATTCCTGATTCATACAGATTAACGCTTATTTGCAATA 1222
Qy 372 TGAGGGGATGTTAATGTCACCAACAAGCACAATTGAGCACTGGAATCCGTGAGCTTA 431
Db 1223 TGATGGGATGTAAAGGCGCTTACCTAACGCAATTCGCAATTTGGAATCGGTGAA 1282
Qy 432 TTTGCCGATATACAGATATGATCCCGAGCTTTTCACAAAGGGGCACTAGAGCTTGA 491
Db 1283 TTTACCTTATATATAATAGACCGCAATTTATTTCCAGAGAGCTTTCTGAATTAATGA 1342
Qy 492 AGCAGTGGCGCTTGATACAAAGCAAAAATAAAGGTGCAATTAAGGCACTGCAACA 551
Db 1343 AACAGTAGCTTTAGATTAATAAGCAAAAGATTAAGAGTAAAGTAAAGATTTGGAACA 1402
Qy 552 AATCGCAAAATTTGGCAATTAAGCAATATGCTATATTTACGCAAGAGCTTGAATTA 611
Db 1403 AATTTGCCAATATGCACAAAATATGATATTAATGTAATCCCAAGCTCAATATGCA 1462
Qy 612 GGTGATGAATGATGTTGCGGCTGAATTTGCAAAAGCGATGTTGCTAGAGCACTACG 671
Db 1463 AGTTTGAATGAGCTGCGCCCTGCTGATTTGAAAAGCAAGCTGCAAAATTAATTTGG 1522
Qy 672 GTTGTATGACAAAGACAGATCGTACGCTTGCATACAGGCTTGTATCAGGTGCA 731
Db 1523 CTATATGACAAAGACAGATGTTGAGTGTGCTGATCTAGGCTTGTATCAGGAAAGAA 1582
Qy 732 TGAATTTGATGATGAAAGCTTTCCGCGGAAAAATTAATCTGATTTATGATTTGGAG 791
Db 1583 TGAATTTGATGATGAAAGCTTTCCGCGGAAATTAATCTGATTTATGATTTGGAG 791
Qy 792 GACGATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 851
Db 1643 AACGATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1702
Qy 852 AACGCTCTCACTAATTAAGAAATGCGCTCAGGGAATCTAGCTTTCATCTATCAT 911
Db 1703 AATATG---CTACAAATTAAGAAATGCGCTCAGGGAATCTAGCTTTCATCTATCAT 1759
Qy 912 GGTATGCGGTGGGGGACTTGAAGACTTACCTTCAATCTGCAATCTTATTCAGCAAGC 971
Db 1760 GGTATGCGGTGGGGGACTTGAAGACTTACCTTCAATCTTATTCATCTATTCAGCAAGC 1819
Qy 972 ATACGCTGCGGTCGCAAGATTCATCAAACTCTGCGGAGCAAGCAAGTGAATGGGCTTA 1031
Db 1820 ATATGCTGCTGAGCGCAAAATTCATCAAACTCTGCGGAGCAAGCAAGTGAATGGGCTTA 1879
Qy 1032 CACAAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1091
Db 1880 TACGACAGCTCTCGAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1939
Qy 1092 CGCTGCGGGAATGAGACCGAACGCGGGAACATCACTGTCACAGGCAAGCTTAAAA 1151
Db 1940 TCGGCGCGGGAATGAGACCGAACGCGGGAACATCACTGTCACAGGCAAGCTTAAAA 1199
Qy 1152 TCGAATTAAGATGAGACCGAACGCGGGAACATCACTGTCACAGGCAAGCTTAAAA 1211
Db 2000 TCGAATTAAGATGAGACCGAACGCGGGAACATCACTGTCACAGGCAAGCTTAAAA 2059
Qy 1212 TATCAACATGTCGACAGTTCTTCTTCACTGTCACAGGCAAGGATGATGATCAAC 1271
Db 2060 TATTAACATGTTGCTCAATTTCTTCACTGTCACAGGCTCTTCACTGATGATGATTAAGCC 2119

Qy 1272 GATGTCATGCGACCGGGAACGTTCACTACTATGACGAAGTCTTCTTTCACCGATTC 1331
Db 2120 GAGCTCATGCGACCGAGTACGATATGCTCTGCTAGATCACTTACCTCAGATTC 2179
Qy 1332 CTCCTTCTGGCGCAACCATGACAGTAAATATGATGATGATGATGATGATGATGATGATG 1391
Db 2180 CTGATTTCTGGCAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2239
Qy 1392 ACCGATGCTTCTGGAACGTCGACAGCTTCTGTCATTTTGTGAAAAACAGAGCAT 1451
Db 2240 TCCAAATTTGACAGTAAATGTCACAAATTAAGGAGCAATTTGTGAAAAATAGAGGGT 2299
Qy 1452 CACACCAAGCCCTTCTCTATTAATAACCGCACTGATTTCCGTCGAGCTGACATCGGCT 1511
Db 2300 AACTCTTAAGCTTCTCTTCTTAAACCTGCTTAAATTTGACAGGTCGCGGATGTCGACT 2359
Qy 1512 TGCTACCGGAACGCTAACCAAGATGCGGACAGTGAACATTTGATTAATCCCTGAACGT 1571
Db 2360 TGCTTTCCAAATGCTAACCAAGATGCGGAGAGTACGTTAGATTAATCCCTGAATGT 2419
Qy 1572 TGCTATGTAAGAGTCCAGTTCTCTATCCACCAAGCAAAAGCAAGTCTGTTTAC 1631
Db 2420 CGCATTTGTAATGAAACGAGCCCTTATCAACAAGTCAAAAGCAACATATTCGTTTAC 2479
Qy 1632 TGCTATGCGCGGACGCTTGAATAATCTCCCTGATATGCTGATGCTCCGTCGAGCAC 1691
Db 2480 GGTCAAGCTGTGTAACCTCTTAAATAATCACTGTTTGTGTCAGATGACCAAGTGGCAC 2539
Qy 1692 AACTGCTCGTACGCTTGTCAATGATCTGACCTTGTCAATACGCTCAAAATGGCAC 1751
Db 2540 GACGCACTACCTTCTTGTGATATTTAGCTTATGATATCTGACCAATGGAAC 2599
Qy 1752 ACAGATGTAAGAAATGACTTTTCTGCGCATACATGATTAACCTGGAATGCGCAATTA 1811
Db 2600 TAAATAGCTCGAAATGACTTTTACAGCACCGTATGATTAACAATTTGGATGCGAAGAAC 2659
Qy 1812 CTTAGAAATGATTTTATTAATGACCAACAAAGCGGAGTATTAATGAGTATGAGG 1871
Db 2660 CTTGAAATGATTTTATTAATGATGCTCTCAAAAGCGGAGTATTAACGATGATGAGG 2719
Qy 1872 TTATAAGTACCGGTTGACCAAGACCTTCTGTTGGCAATTTGAATTA 1923
Db 2720 TTACATGTAAGTAAAGTCCGCAACCTTTCTTGAAGATTTGACATTA 2771

RESULT 8
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, P.
APPLICANT: SCHIEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```
/ APPLICATION NUMBER: US/07/935,313
/ FILING DATE:
/ APPLICATION NUMBER: EP 91 114 300-6
/ FILING DATE: 26-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/114 IMMU
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-9300
/ TELEFAX: (703) 683-4109
/ TELEX: 899149
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7218 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ CLONE: pTZapc-P18
/ US-08-232-463-14

Query Match 1%, 2.8%, Score 54; DB 2; Length 7218;
Best Local Similarity 2.4%, Pred. No. 8.5e-06;
Matches 9; Conservative 222; Mismatches 147; Indels 0; Gaps 0;

QY 468 AAAAGGCGCATGAGCTGTAAAGACATGCGCTTGTATACAAAGCAAAATTAAGA 527
DB 1430 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1371
QY 528 GGTGCAATTAAAGGACATGACAAATTCGACAAATTCGCAATTAAGATGCTGTATA 587
DB 1370 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1311
QY 588 TATTACGGAAGCGTGAATGATGATGATGATGATGATGATGATGATGATGATG 647
DB 1310 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1251
QY 648 GGATGTGCTCAGACAGCATGACGGTTGTATGACAAAGACAGATGCGGTGCCGA 707
DB 1250 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1191
QY 708 TACAGGCTTGTATGACGTCGCAATGACACTGATGATGATGATGATGATGATG 767
DB 1190 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1131
QY 768 TACTGATTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 827
DB 1130 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1071
QY 828 GCATGTGCTGCTGCTGCT 845
DB 1070 RRRRATCGCAAGCTCCT 1053

RESULT 9
US-09-621-976-2574
/ Sequence 2574, Application US/09621976
/ Patent No. 6639063
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J. B.
/ APPLICANT: Jobert, S.
/ APPLICANT: Giordano, J. Y.
/ TITLE OF INVENTION: ESTs and Encoded Human Proteins.
/ FILE REFERENCE: GENSER.054PR2
/ CURRENT APPLICATION NUMBER: US/09/621,976
/ CURRENT FILING DATE: 2000-07-21
/ NUMBER OF SEQ ID NOS: 19335
/ SOFTWARE: Patent.pm
/ SEQ ID NO 2574
/ LENGTH: 834
/ TYPE: DNA
/ ORGANISM: Homo sapiens
```

```
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 497..826
/ US-09-621-976-2574

Query Match 2.2%, Score 42.4; DB 3; Length 834;
Best Local Similarity 16.9%, Pred. No. 0.012;
Matches 28; Conservative 81; Mismatches 57; Indels 0; Gaps 0;

QY 822 TGTAAGCATGTGCTGCTCGTATTAAGAAACGCTCCATTAATAAGATGCGGCC 881
DB 38 WGSCTSTBRMRSKSYGRACMSRRGRKACMSWTGMSCTGRMSWKRKRYSWYSTK 97
QY 882 TCAGCGAATTAATCTTCCAAATCTATCATCATGATGAGGAGCTTGAGAGACTAC 941
DB 98 WKMRGYSWTYSYRMYRCYSMMWMCMSCTKSRSMCMGSGRSMCMGTGMSAARCRVC 157
QY 942 TTCGATCTGCAAACTTATTCAGCAAGCATATGATGCTGTGCC 987
DB 158 MRYRMTCAKSRKMKMKMKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMS 203

RESULT 10
US-09-620-312D-590/C
/ Sequence 590, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyun
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yungling
/ APPLICANT: Wang, Duntui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ FILE REFERENCE: 784C1P2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ CURRENT FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 590
/ LENGTH: 5053
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (154) ..(2409)
/ US-09-620-312D-590

Query Match 2.2%, Score 41.8; DB 3; Length 5053;
Best Local Similarity 57.1%, Pred. No. 0.056;
Matches 76; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 174 TGTGCTTTTGTGTTGATGTAATAATGTGAATCCCAAAAGTTTGCAAGAGCT 233
DB 3158 TGAAGTTATGTGATGTAACCAAGCTACTGAAACCTGCAATGCTGTTTAAAGTTT 3099
QY 234 TGAACAGTCCCGCAATTAATTAATTCATATTCATTAATTCATTAATTCATTA 293
```



```
Db      3098 TCTACCGTACACACTACTACTACCTCCATACCAACCAACTTTGTCTCAATTAACA 3039
Qy      294 AGAACAACAAACAG 306
        |||
Db      3038 TGAGACAAACAG 3026

RESULT 11
US-09-949-016-16621/c
; Sequence 16621, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16621
; LENGTH: 108341
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16621

Query Match      2.2%; Score 41.4; DB 3; Length 108341;
Best Local Similarity 55.1%; Pred. No. 0.51;
Matches 81; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy      380 ATGTTAAGTCAGACACAGACCATTTGAGACGTCGATCCGTCGACCTTATTTGCCGA 439
        |||
Db      103396 ATGTTAAGTCAGACACAAACAAATTTGAGATTTCAATTAAGTAATTTATTTGCCCA 103337
        |||
Qy      440 TATACAGATATGATCCCAAGCTTTTCACAAAGGGGCGATCAGCTTGTAAAGCAGTGG 499
        |||
Db      103336 TAAAGAGATGAAGTATGATTCATGCTACATGATGAATCTTGAAGACATGTGC 103277
        |||
Qy      500 CGCTGATACAAAGCAGAAATTAAG 526
        |||
Db      103276 TAAATGAAGAGGACAGACCAAAAG 103250

RESULT 12
US-09-949-016-16601/c
; Sequence 16601, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16601
; LENGTH: 90724
; TYPE: DNA
; ORGANISM: Human
```

```
US-09-949-016-16601
Query Match      2.1%; Score 39.8; DB 3; Length 90724;
Best Local Similarity 49.3%; Pred. No. 1.5;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

Qy      111 TTTCAGGAATTCACAGCAACATCTGATGCTAAAGTTTCTCAACAGCGGCGAGCTGCG 170
        |||
Db      39491 TATTATTGGAAGAGACAGAAAGAAATGAGCAGGGGTGCTATGAAATTAACCTGTATTT 39432
        |||
Qy      171 TGCTGCTGCTTTTCTGTGGAATCTGAAATGTAAGTCCCAAGTTTGCAGAGAA 230
        |||
Db      39431 TGAATGGAGATTGAGTGAATTAATCAAGATTAACTTAATTTTAAGTGAAGTCAAGAA 39372
        |||
Qy      231 GCTTGAACAGTCCCGCAATATTAATCTCATATTAATCAATTCAGACCAATTTT 290
        |||
Db      39371 GCCCTACACAGCCCTCGAAATCTAAAGAAAGTCTCTCAATCATCAGGCCAAAGCA 39312
        |||
Qy      291 AGAAGAAACAAACAGCAGCTGAAAAACA 321
        |||
Db      39311 AGCTAATCTAAACAAACATGATGGCAACA 39281
        |||

RESULT 13
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.c., L.a., and B.n. PAB1 promoters
US-09-806-708B-22

Query Match      2.0%; Score 38.8; DB 3; Length 1141;
Best Local Similarity 9.6%; Pred. No. 0.2;
Matches 60; Conservative 253; Mismatches 309; Indels 1; Gaps 1;

Qy      2 TGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTCAGCGATTTTGTGACTG 61
        |||
Db      691 TDARRNNNTTVARRRMTTKTRWSTTRRHHTGATNNNNNNNNNNNNNNNNNNNNNNNNNNNN 632
        |||
Qy      62 TTGCGTTAAGTATTCATCTGCAAGTGTGCAAGAAATTTGATCTGATTTCAAGAA 121
        |||
Db      631 MMTMTWTMGDGTATKVKWMDTTCYDVADSVWVYAMWRCRDVYTRNNNTYCKS 572
        |||
Qy      122 TTCAGCAACAACTGATGTAAAGTTTCTCAAGCAGGGGCGAGCTGTGCTGCTT 181
        |||
Db      571 YAHASYWWSNNMMWRRYSARWWSVAAWTTRNNMMWSGVRRWAGTMMWRMNNNNNTD 512
        |||
Qy      182 TTCTGTGGAATCTGAAAATGTGAACCTCCCAAGAGTTTGCAGAGAAAGTGAAGACAG 241
        |||
Db      511 TTYVMMKXKABBTIVYDSMCNAKSMRGNMRRMAMKMMWANNNDAGADHHTYMMGNNT 452
        |||
Qy      242 TCCCGGCAATTAATACTCAATATTCAATTCAGAGCAACATTTAGAAAGAAACA 301
        |||
Db      451 MMRRAVMMKMMNAWCBAYCCNNNNNNRACVWHKHKMWRTYKMKWZACNNNNBKAMVNR 392
        |||
Qy      302 AACAGAGCTGAGAAAAAAGAGGGGCAAGATTCTGACATCAATCTGATTAATGCTTAA 361
        |||
```


Db 391 VAMMMYSRDTNTDMMWMTSDMBWHYTVDYTMWRAANNNNNNNNRBCKTSSMMMMMDHM 332
Qy 362 TTGTGAGATATAGGCGCATGTTAAGTCGCAACAGCAGCATGAGCGATGCG 421
Db 331 NTHCTGNNNTWGSABMAMSMWAGASBYTYNNCMRTYNGKTMNNNNNK-AMYR 273
Qy 422 TGGACCTTATTTGCGATATACAGAAATGATCCCGCATTTTCAAAAAGGCGATCG 481
Db 272 TKTAVACNNRRYYTDAVTVBKNYKCYBYBYTYBYTKHMBWRBAPHNSMMWYK 213
Qy 482 AGCTTGTAAGACGATGCGGCTTGATACAAACAGAAAATAAGAGTGCAATTAAAG 541
Db 212 CNKTYVSNYHYAMBYKBAVGVGNNNNKDRMHHNMCAITNNNNMMWYAYVHHMKKG 153
Qy 542 GCGTGAACAAATCCGCAATTCGCAATAGCAATGATGCTATATATTAAGCGCAACG 601
Db 152 KAAWNNNTKABRDDHBAHVKTYYWRYDYWCAMCMMNAKAVRTAKMHHYTTDRVSN 93
Qy 602 CTAGATATAGGATGATGAT 624
Db 92 NTGVAMMMRRCMMWYSMMNRY 70

RESULT 14
US-08-750-532-2
; Sequence 2, Application US/08750532
; Patent No. 5756339
; GENERAL INFORMATION:
; APPLICANT: MITTA, Masanori
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MORISHITA, Mio
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,532
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01095
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 1994/130236
; FILING DATE: 13-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 1994/173912
; FILING DATE: 26-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2835 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-750-532-2
Query Match 2.0%; Score 37.6; DB 2; Length 2835;
Best Local Similarity 54.3%; Pred. No. 0.86;
Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
Qy 1037 CAGATTCAGAAATGTGATGATGATGCGGCAAAAATGATATGACATCTTTGCTG 1096
Db 1529 CAGAAAGGTTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1588
Qy 1097 CCGGATGAGAGACCGAAGCGGGAACCATCATGACGACGACGACGACGACGACG 1156
Db 1589 CAGGAATGAAAGCTCTGCGCATTAACATGTTGGAAGTCTGCTGTTGCAACAAAGCG 1648
Qy 1157 TAAAGTCGAGCTACGGA 1176
Db 1649 TAACTGTGAGAGCTGCTCA 1668

RESULT 15
US-08-750-532-8
; Sequence 8, Application US/08750532
; Patent No. 5756339
; GENERAL INFORMATION:
; APPLICANT: MITTA, Masanori
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MORISHITA, Mio
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,532
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01095
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 1994/130236
; FILING DATE: 13-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 1994/173912
; FILING DATE: 26-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4765 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: April 8, 2006, 01:34:45 ; Search time 1380.58 Seconds
(without alignments)
11518.383 Million cell updates/sec

Title: US-10-784-870-5

Perfect score: 1923
Sequence: 1 atgagagaagagagaaagct.....cgttgcaatctgtaattaa 1923

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main:*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
8: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1923	100.0	1923	3	US-09-920-954-5
2	1923	100.0	1923	7	US-10-456-479-3
3	1923	100.0	1923	7	US-10-784-870-5
4	1923	100.0	1923	8	US-10-820-712A-2
5	1923	100.0	1923	9	US-10-820-714A-2
6	1923	99.8	1923	3	US-09-920-954-7
7	1919.8	99.8	1923	7	US-10-784-870-7
8	1477.2	76.8	1920	3	US-09-920-954-3
9	1477.2	76.8	1920	7	US-10-784-870-3
10	1305	67.9	1305	6	US-10-385-662-1
11	1281.8	66.7	1302	8	US-10-820-712A-13
12	1281.8	66.7	1302	8	US-10-820-714A-14
13	997	51.8	1302	8	US-10-820-712A-11
14	997	51.8	1302	9	US-10-820-714A-12
15	854.6	44.4	1302	8	US-10-820-712A-21
16	854.6	44.4	1302	8	US-10-820-714A-22
17	740.4	38.5	1299	8	US-10-820-712A-19
18	740.4	38.5	1299	8	US-10-820-714A-20
19	738.8	38.4	1299	8	US-10-820-712A-15
20	738.8	38.4	1299	8	US-10-820-714A-16
21	737.2	38.3	1299	8	US-10-820-712A-17
22	737.2	38.3	1299	9	US-10-820-714A-18
23	48	2.5	54	6	US-10-385-662-20

C	24	45	2.3	415	5	US-10-016-349A-101	Sequence 101, App
C	25	41.8	2.2	1095	5	US-10-106-698-1076	Sequence 1076, App
C	26	41.8	2.2	1297	3	US-09-822-846-311	Sequence 311, App
C	27	41.8	2.2	4440	9	US-10-756-149-4271	Sequence 4271, App
C	28	41.8	2.2	5053	5	US-10-037-270-590	Sequence 590, App
C	29	41.8	2.2	5053	6	US-10-117-722-590	Sequence 590, App
C	30	41.8	2.2	5053	9	US-10-122-851-590	Sequence 590, App
C	31	41.8	2.2	8195	3	US-09-764-891-9013	Sequence 9013, App
C	32	40.2	2.1	604	4	US-09-925-065A-3789	Sequence 3789, App
C	33	40	2.1	834	5	US-10-123-155-148	Sequence 148, App
C	34	40	2.1	834	6	US-10-146-731-148	Sequence 148, App
C	35	40	2.1	834	6	US-10-140-672-148	Sequence 148, App
C	36	40	2.1	834	6	US-10-141-761-148	Sequence 148, App
C	37	40	2.1	834	6	US-10-142-885-148	Sequence 148, App
C	38	40	2.1	834	6	US-10-158-790-148	Sequence 148, App
C	39	40	2.1	834	6	US-10-137-871-148	Sequence 148, App
C	40	40	2.1	834	6	US-10-140-923-148	Sequence 148, App
C	41	40	2.1	834	6	US-10-141-756-148	Sequence 148, App
C	42	40	2.1	834	6	US-10-141-759-148	Sequence 148, App
C	43	40	2.1	834	6	US-10-140-805-148	Sequence 148, App
C	44	40	2.1	834	6	US-10-140-864-148	Sequence 148, App
C	45	40	2.1	1330	9	US-10-956-157-281	Sequence 281, App

ALIGNMENTS

RESULT 1

US-09-920-954-5
Sequence 5, Application US/09920954
Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, MATSUHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920, 954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509, 814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-09-920-954-5

Query Match 100.0%; Score 1923; DB 3; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCGACGACGATTTGTGACT	60
DB	1	ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCGACGACGATTTGTGACT	60
QY	61	GTTCCTTAAGTATCATCTGCAAGTGTGCAAGAAATTTGATCTGATTTCAAGA	120
DB	61	GTTCCTTAAGTATCATCTGCAAGTGTGCAAGAAATTTGATCTGATTTCAAGA	120

SEQ ID NO 3
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp. KSM-KP43
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
OTHER INFORMATION:
US-10-456-479-3

Query Match 100.0%; Score 1923; DB 7; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAAAGAAAGAGTCTTTTATCTGTTTATCAAGCTGACGCAATTTGTCACT 60
DB 1 ATGAGAAAGAAAGAGTCTTTTATCTGTTTATCAAGCTGACGCAATTTGTCACT 60
QY 61 GTTGCCTTAAGTAATCCATCTGACGTGTGCAAGAAATTTGATCTGAAATTTCAAGGA 120
DB 61 GTTGCCTTAAGTAATCCATCTGACGTGTGCAAGAAATTTGATCTGAAATTTCAAGGA 120
QY 121 ATTCAACAACTGATGCTTAAAGTTTCTCCAGCAGGGGCAAGCTGTGCTGCT 180
DB 121 ATTCAACAACTGATGCTTAAAGTTTCTCCAGCAGGGGCAAGCTGTGCTGCT 180
QY 181 TTTCTGTGGAATCTGAAAATGTGAACTCCCAAAAGTTTGCAGAAAGCTTGAACA 240
DB 181 TTTCTGTGGAATCTGAAAATGTGAACTCCCAAAAGTTTGCAGAAAGCTTGAACA 240
QY 241 GTCCCGCAATATAATCCATATTCATTCATTCAGTGAACCAATTTTGAAGAAACA 300
DB 241 GTCCCGCAATATAATCCATATTCATTCATTCAGTGAACCAATTTTGAAGAAACA 300
QY 301 AAAACAGAGCTGAAAAAACAAGGCGCAAGATTTCTGACTACATACCTGATTAAGCTTAC 360
DB 301 AAAACAGAGCTGAAAAAACAAGGCGCAAGATTTCTGACTACATACCTGATTAAGCTTAC 360
QY 361 ATTGTGAGATGAGGGGCGATTTTAACTGACCAACAGCAACCATTTAGCAAGCTGGAATCC 420
DB 361 ATTGTGAGATGAGGGGCGATTTTAACTGACCAACAGCAACCATTTAGCAAGCTGGAATCC 420
QY 421 GTGAGGCTTAATTTGCGATATACAGATAGATCCCGAGCTTTTCAAAAAGGCGATCA 480
DB 421 GTGAGGCTTAATTTGCGATATACAGATAGATCCCGAGCTTTTCAAAAAGGCGCGATCA 480
QY 481 GAGCTTGTAAACAGTGGCGCTTGAATACAAAGCAAGAAATTAAGAGTGCATTTAAGA 540
DB 481 GAGCTTGTAAACAGTGGCGCTTGAATACAAAGCAAGAAATTAAGAGTGCATTTAAGA 540
QY 541 GGCATGGAACAAATCGCACAATTCGCAATAGCAATGATGTCTATATTTACGGCAAG 600
DB 541 GGCATGGAACAAATCGCACAATTCGCAATAGCAATGATGTCTATATTTACGGCAAG 600
QY 601 CCGAATATTAAGGTGATGAAATGATTTGCGCGGTGGAATTTGTCAAAGGAGTGTGCTCAG 660
DB 601 CCGAATATTAAGGTGATGAAATGATTTGCGCGGTGGAATTTGTCAAAGGAGTGTGCTCAG 660
QY 661 AGCAGCTACGGGTTGTATGACAAAGACAGATGTAAGCGGTGCGCATACAGGGCTTGAT 720
DB 661 AGCAGCTACGGGTTGTATGACAAAGACAGATGTAAGCGGTGCGCATACAGGGCTTGAT 720
QY 721 ACAGGTGCAATGACATTCGATGCAATGAAGCTTCGCGGGAATAATTAATCTGCAATTAAT 780
DB 721 ACAGGTGCAATGACATTCGATGCAATGAAGCTTCGCGGGAATAATTAATCTGCAATTAAT 780
QY 781 GCAATGGGACGGAAGATTAATGCAATGATAGCAATGATGATGATGATGATGATGATGAT 840
DB 781 GCAATGGGACGGAAGATTAATGCAATGATAGCAATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGTAATAGGAACGGCTCACTAATTAAGGAATGGCGCTCAGGCGAATCTAGCTTTC 900
DB 841 TCCGTAATAGGAACGGCTCACTAATTAAGGAATGGCGCTCAGGCGAATCTAGCTTTC 900

QY 901 CAATCTATCATGATTAACGGGTGGGGAATTTGAGAGACTACCTTGAATCTGCAAACTTTA 960
DB 901 CAATCTATCATGATTAACGGGTGGGGAATTTGAGAGACTACCTTGAATCTGCAAACTTTA 960
QY 961 TTGAGCGAAGATACAGTGTGCGGCAAGATTTCAATACTCTGAGGAGCAAGCTG 1020
DB 961 TTGAGCGAAGATACAGTGTGCGGCAAGATTTCAATACTCTGAGGAGCAAGCTG 1020
QY 1021 AATGGGCTTACACACAGATTCAGAAATGTGATGATATGTCGCAAAATGATATG 1080
DB 1021 AATGGGCTTACACACAGATTCAGAAATGTGATGATATGTCGCAAAATGATATG 1080
QY 1081 ACGATCTTTTCTGCTGCGGGAATGAAGAACCGAAGCGCGAATCATGATGACCAAGC 1140
DB 1081 ACGATCTTTTCTGCTGCGGGAATGAAGAACCGAAGCGCGAATCATGATGACCAAGC 1140
QY 1141 ACGCTAAATAATCAATTAAGTCCGAGCTAACGGAATCTCCGCGCAAGCTTGGGCT 1200
DB 1141 ACGCTAAATAATCAATTAAGTCCGAGCTAACGGAATCTCCGCGCAAGCTTGGGCT 1200
QY 1201 TATGCGCAATATCAACATGATGCAAGTCTCTTCAAGTGAACGCAAGAGATGA 1260
DB 1201 TATGCGCAATATCAACATGATGCAAGTCTCTTCAAGTGAACGCAAGAGATGA 1260
QY 1261 CGGATCAACCGGATGTCATGCAACGGGAACTTCACTATACGAAGATCTTCTT 1320
DB 1261 CGGATCAACCGGATGTCATGCAACGGGAACTTCACTATACGAAGATCTTCTT 1320
QY 1321 GCAACGGATCTCTCTTCTGCGGGAACAGTAAAGTAAATGATCAATGCGTGAACG 1380
DB 1321 GCAACGGATCTCTCTTCTGCGGGAACAGTAAAGTAAATGATCAATGCGTGAACG 1380
QY 1381 TCCATGCTACACCGATCTGCTGGAACAGTGAACAAGCTTGTGAGCAATTTGTGAAA 1440
DB 1381 TCCATGCTACACCGATCTGCTGGAACAGTGAACAAGCTTGTGAGCAATTTGTGAAA 1440
QY 1441 AACAGAGCATCAACCAAAAGCTTCTATTAATAAGCGCACTGATTCGGTGAAGCT 1500
DB 1441 AACAGAGCATCAACCAAAAGCTTCTATTAATAAGCGCACTGATTCGGTGAAGCT 1500
QY 1501 GACATGCGCTTGGCTAACCGGAACGGTAACCAAGATGAGGGAAGATGATGATTA 1560
DB 1501 GACATGCGCTTGGCTAACCGGAACGGTAACCAAGATGAGGGAAGATGATGATTA 1560
QY 1561 TCCCTGAACGTTGCTATGTAAGAGTCAAGTCTCTATTCACAGCCAAAGAGAGC 1620
DB 1561 TCCCTGAACGTTGCTATGTAAGAGTCAAGTCTCTATTCACAGCCAAAGAGAGC 1620
QY 1621 TACTGCTTAATGCTTACCTGCGGCAAGCTTTGAAAATCTCCCTGATATGATGATGCC 1680
DB 1621 TACTGCTTAATGCTTACCTGCGGCAAGCTTTGAAAATCTCCCTGATATGATGATGCC 1680
QY 1681 CCGCGAGCAACATGCTTCCGTAAGCTTTGTGAATGTGGAACCTTGTCAATTAAGCT 1740
DB 1681 CCGCGAGCAACATGCTTCCGTAAGCTTTGTGAATGTGGAACCTTGTCAATTAAGCT 1740
QY 1741 CCAATGGCAACAGATATGAGAAATGACTTACTTCCGCAATACATGATTAATCTGAGAT 1800
DB 1741 CCAATGGCAACAGATATGAGAAATGACTTACTTCCGCAATACATGATTAATCTGAGAT 1800
QY 1801 GGCAGCAATTAAGTAAATTAATTAATGACCAAAAGCGGAGCTATTAACATTT 1860
DB 1801 GGCAGCAATTAAGTAAATTAATTAATGACCAAAAGCGGAGCTATTAACATTT 1860
QY 1861 GAGGTACAGCTTATTAAGTACCGTTGGAACCAAGCTTCTGCTGGAATTTGGAAT 1920
DB 1861 GAGGTACAGCTTATTAAGTACCGTTGGAACCAAGCTTCTGCTGGAATTTGGAAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923

RESULT 3

US-10-784-870-5

: Sequence 5, Application US/10784870

Publication No. US20040142837A1

; GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO

APPLICANT: OKUDA, MITSUYOSH

APPLICANT: SAEKI, KATSUHI

APPLICANT: KUBOTA, HIROMI

APPLICANT: HITOMI, JUN

APPLICANT: KAGEYAMA, YASUSHI

APPLICANT: SHIKATA, SHITSUMI
NOMINER: MASAEYUMI

APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE

FILE REFERENCE: 0327-0832-0PCT

! FILE REFERENCE: 032/-0632-0PCI
! CURRENT APPLICATION NUMBER: IIS/10/784.870

CURRENT FILING DATE: 2004-02-24

CONSENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509.814A

PRIOR FILING DATE: 2000-04-06

PCT/JP98/04528

PRIOR FILING DATE: 1998-10-07 ;

PRIOR APPLICATION NUMBER: JP 9-274570

;
PRIOR FILING DATE: 1997-06-08

; NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn version 3.0.0

; SEQ ID NO 5

; LENGTH: 1923

TYPE: DNA

ORGANISM: *Bacillus* sp.

NAME/KEY:

NAME/KEY: CDS
LOCATION: (1) (1923)

LOCATIONS: (1) .. (1943)
US-10-784-870-5

05-10704-00-2

Query Match	100.0%	Score 1923:	DB 7:	Length 1923:
-------------	--------	-------------	-------	--------------

Best Local Similarity	100.0%	Pred. No. 0:
2007.0000	100.00%	Score 1525/

Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY	1	ATGAGAAAAGAAAAAGGTGTTTATCTGTTTATCAGCTGCAGCAATTTGTGCAC	60
Db	1	ATGAGAAAAGAAAAAGGTGTTTATCTGTTTATCAGCTGCAGCAATTTGTGCAC	60
QY	61	GTTGCGTTAAGTAAATCCATCTGCAGTGGTGCAGAAATTTTGATCTGGAATTTCAAGGA	120
Db	61	GTTGCGTTAAGTAAATCCATCTGCAGTGGTGCAGAAATTTTGATCTGGAATTTCAAGGA	120
QY	121	ATTCAAGCAACAACATGATGCTAAAGGTTTCTCAAGCAGGGGCAAGCTGAGTCTGCT	180
Db	121	ATTCAAGCAACAACATGATGCTAAAGGTTTCTCAAGCAGGGGCAAGCTGAGTCTGCT	180
QY	181	TTTCTGTGGAAATCTGAAAAATGTGAATCTCCCAAAAGGTTTGCAGAGAAAGCTTGAACA	240
Db	181	TTTCTGTGGAAATCTGAAAAATGTGAATCTCCCAAAAGGTTTGCAGAGAAAGCTTGAACA	240
QY	241	GTCGCGGCAAAATTAATACTCATATTTATCCAAATTCATATGACCAATTTTGAAGAAACA	300
Db	241	GTCGCGGCAAAATTAATACTCATATTTATCCAAATTCATATGACCAATTTTGAAGAAACA	300
QY	301	AAACAGCAGCTGGAAAAAACAAGGCAAAAGATCTTCGACTACATACCTGATTTATGCTTAC	360
Db	301	AAACAGCAGCTGGAAAAAACAAGGCAAAAGATCTTCGACTACATACCTGATTTATGCTTAC	360
QY	361	ATTGTGCAATTAAGAGGGCGAATGTTAAATCAGCAACAGACCAATTTGAGCAGTGGAAATCC	420
Db	361	ATTGTGCAATTAAGAGGGCGAATGTTAAATCAGCAACAGACCAATTTGAGCAGTGGAAATCC	420
QY	421	GTCGAGCTTATTTGCGCATATACGAATATAGATCCCGAGCTTTTCACAAAAAGGGGCATCA	480
Db	421	GTCGAGCTTATTTGCGCATATACGAATATAGATCCCGAGCTTTTCACAAAAAGGGGCATCA	480
QY	481	GAGCTTGTAAAAAGCAGTGGCGCTTGATACAAAGCAAAAAATAAAGAGGTGCATTTAAGA	540

Db	481	GAGCTGTAAAAAGCAATGCGCCTTGATATCAAAAGCAAAAAATTAAGAGGTGCATTAA	540
Qy	541	GGCATCGAA CAAATCGCAAAATTCGCATTAAGCAATGATGTCTATATATTA CCGCAAG	600
Db	541	GGCATCGAA CAAATCGCAAAATTCGCATTAAGCAATGATGTCTATATATTA CCGCAAG	600
Qy	601	CCGAGCTAATAGGTGA TGAATGTGTGGCGGTGAAATTTGTCAAAGCGGATGTGGTCA	660
Db	601	CCGAGCTAATAGGTGA TGAATGTGTGGCGGTGAAATTTGTCAAAGCGGATGTGGTCA	660
Qy	661	AGCAGCTACGGGTGTGATGAGCAAGGACAGATCTGACGGTTCGCAATACAGGCTTGAT	720
Db	661	AGCAGCTACGGGTGTGATGAGCAAGGACAGATCTGACGGTTCGCAATACAGGCTTGAT	720
Qy	721	ACAGGTCGCAATGACAGTTTCGATGCA TGAAGCTTTCGCGGAAAAATTA CTCGATTATAT	780
Db	721	ACAGGTCGCAATGACAGTTTCGATGCA TGAAGCTTTCGCGGAAAAATTA CTCGATTATAT	780
Qy	781	GCATTGGGACGGACGGAATTAATGCCAATGATATAGAAATGGTCATGGTACGATGGCTGGC	840
Db	781	GCATTGGGACGGACGGAATTAATGCCAATGATATAGAAATGGTCATGGTACGATGGCTGGC	840
Qy	841	TCGGTATTAAGAAACGGCTTCA CTATATAAAGGAATGGCGCTCAGGCGAATCTAGCTTTC	900
Db	841	TCGGTATTAAGAAACGGCTTCA CTATATAAAGGAATGGCGCTCAGGCGAATCTAGCTTTC	900
Qy	901	CAATCTATCATGGAATTAAGCGGTGGGGGACCTTGGAGGACATACCTTCGAAATCTGCAACCTTA	960
Db	901	CAATCTATCATGGAATTAAGCGGTGGGGGACCTTGGAGGACATACCTTCGAAATCTGCAACCTTA	960
Qy	961	TTCAAGCCAAAGCATACAGTGTGTGTGTC CAGAAATTCATACAAACTCTCGGGGAGCAGCAGTG	1020
Db	961	TTCAAGCCAAAGCATACAGTGTGTGTGTC CAGAAATTCATACAAACTCTCGGGGAGCAGCAGTG	1020
Qy	1021	AATGGGGCTTACACAA CAGATTC CAGAAATGTGATGACTATGTGCGCAAAAATGATATG	1080
Db	1021	AATGGGGCTTACACAA CAGATTC CAGAAATGTGATGACTATGTGCGCAAAAATGATATG	1080
Qy	1081	ACGATCTCTTTCGCTGCGCGGGA TGAAGA CCGAACGGCGGAA CCAATCAGTGCACAGGC	1140
Db	1081	ACGATCTCTTTCGCTGCGCGGGA TGAAGA CCGAACGGCGGAA CCAATCAGTGCACAGGC	1140
Qy	1141	ACAGCTAAAAATGCATATACAGTCCGAGCTACGAAAA CTTCCGCCCAAGCTTTGGGTCT	1200
Db	1141	ACAGCTAAAAATGCATATACAGTCCGAGCTACGAAAA CTTCCGCCCAAGCTTTGGGTCT	1200
Qy	1201	TATGGGCAATATATCAACATATGTGGCAAGTCTTTCA CGTGGACCGACAAAGATGGA	1260
Db	1201	TATGGGCAATATATCAACATATGTGGCAAGTCTTTCA CGTGGACCGACAAAGATGGA	1260
Qy	1261	CGGATCAAAACGGATGTGATGGCACCGGGAAAGTTTCACTATCAGCAAGATCTTCTCTT	1320
Db	1261	CGGATCAAAACGGATGTGATGGCACCGGGAAAGTTTCACTATCAGCAAGATCTTCTCTT	1320
Qy	1321	GCACTCGGATTCCTCTTCTTGGGCGAACATGACAGTAAATATATGCAATATGCGGTGGAAACG	1380
Db	1321	GCACTCGGATTCCTCTTCTTGGGCGAACATGACAGTAAATATATGCAATATGCGGTGGAAACG	1380
Qy	1381	TCGATGGCTACACCGATCGTTGTGCGAAACGTTGGGACAGCTTCGAGACATTTTGTGAAA	1440
Db	1381	TCGATGGCTACACCGATCGTTGTGCGAAACGTTGGGACAGCTTCGAGACATTTTGTGAAA	1440
Qy	1441	AAACAGGCGATCAACCAAGCTTCTCTATTTAAAGCGGACATGATTCGCGGTGACCT	1500
Db	1441	AAACAGGCGATCAACCAAGCTTCTCTATTTAAAGCGGACATGATTCGCGGTGACCT	1500
Qy	1501	GACATCGGCTTGGCTATCCGAAACGGTATCCAAAGATGGGGACGAGTGACATTTGGATAAA	1560
Db	1501	GACATCGGCTTGGCTATCCGAAACGGTATCCAAAGATGGGGACGAGTGACATTTGGATAAA	1560
Qy	1561	TCCTCGAAGCTTGGCTATGATGAACGAGTCAAGTTCTCTATACAGCGCAAAAAGCGACG	1620
Db	1561	TCCTCGAAGCTTGGCTATGATGAACGAGTCAAGTTCTCTATACAGCGCAAAAAGCGACG	1620

QY 1621 TACTGTTTACTGCTACGCGGCAAGCTTTGAAAACTCCCTGCTATGCTGATGCC 1680
Db 1621 TACTGTTTACTGCTACGCGGCAAGCTTTGAAAACTCCCTGCTATGCTGATGCC 1680
QY 1681 CTTGCGGCAACACTGCTTCCGTTACCGCTTGTCAATGATCTGACCTTGTCAATGCC 1740
Db 1681 CTTGCGGCAACACTGCTTCCGTTACCGCTTGTCAATGATCTGACCTTGTCAATGCC 1740
QY 1741 CCAATGCGACACAGTATGTAAGAAATGACTTTTCTGCGCATCAATGATTAATCGGAT 1800
Db 1741 CCAATGCGACACAGTATGTAAGAAATGACTTTTCTGCGCATCAATGATTAATCGGAT 1800
QY 1801 GCGCGCAATACGTAGAAAAATGATTTATTAATGACCAACAAAGCGGACGTATACAT 1860
Db 1801 GCGCGCAATACGTAGAAAAATGATTTATTAATGACCAACAAAGCGGACGTATACAT 1860
QY 1861 GAGGTACAGGCTTATTAACGTACCGGTTGACCAACACCTTCTGTTGGCAATTTGAA 1920
Db 1861 GAGGTACAGGCTTATTAACGTACCGGTTGACCAACACCTTCTGTTGGCAATTTGAA 1920
QY 1921 TAA 1923
Db 1921 TAA 1923
RESULT 4
US-10-820-712A-2
Sequence 2, Application US/10820712A
Publication No. US20050026804A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuyoshi
APPLICANT: Izawa, Yoshihumi
APPLICANT: Kobayashi, Tohru
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tetsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-USO
CURRENT APPLICATION NUMBER: US/10/820,712A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106708
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp. KSM-KP43
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
NAME/KEY: sig_peptide
LOCATION: (1)..(618)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (619)..()
US-10-820-712A-2
Query Match 100.0%; Score 1923; DB 8; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTCAGGCAATTTGTGACT 60
Db 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTCAGGCAATTTGTGACT 60
QY 61 GTTGCCTTAAGTATCATCTGCAAGGTGTGCAAGAAATTTGATCTGCAATTTCAAGA 120
Db 61 GTTGCCTTAAGTATCATCTGCAAGGTGTGCAAGAAATTTGATCTGCAATTTCAAGA 120
QY 121 ATTACAGCAACACTGATGCTAAAGTTTCTCAAGAGGGGCAAGCTGCTGCTGCT 180

Db 121 ATTACAGCAACACTGATGCTAAAGTTTCTCCAGAGGGGCAAGCTGCTGCTGCT 180
QY 181 TTTCTGTGGAATCTGAAAATGTGAAAATCTCCAAAAGSTTTGCAAGAACTTGAACA 240
Db 181 TTTCTGTGGAATCTGAAAATGTGAAAATCTCCAAAAGSTTTGCAAGAACTTGAACA 240
QY 241 GTCCCGCAATATATAACTCCATAATTAATCAATTTCAATGAGCAATTTTGAAGAA 300
Db 241 GTCCCGCAATATATAACTCCATAATTAATCAATTTCAATGAGCAATTTTGAAGAA 300
QY 301 AAACAGAGCTGAAAAAAGAGGCAAGATTTCTGACATACACTGATTAATGCTTAC 360
Db 301 AAACAGAGCTGAAAAAAGAGGCAAGATTTCTGACATACACTGATTAATGCTTAC 360
QY 361 ATTGTGAGATTAAGAGGCGATTTTAAGTCAAGCAACAGCACTTGAAGCAAGTATCC 420
Db 361 ATTGTGAGATTAAGAGGCGATTTTAAGTCAAGCAACAGCACTTGAAGCAAGTATCC 420
QY 421 GTGAGGCTTAATTTGCGCATATACAGATGATCCCGCTTTTCAAAAAGGGGCAATCA 480
Db 421 GTGAGGCTTAATTTGCGCATATACAGATGATCCCGCTTTTCAAAAAGGGGCAATCA 480
QY 481 GAGCTGTAAAGCAAGTGGCGCTTGATACAAAGCAAAAAATTAAGAGTGCATTAAGA 540
Db 481 GAGCTGTAAAGCAAGTGGCGCTTGATACAAAGCAAAAAATTAAGAGTGCATTAAGA 540
QY 541 GGCATGCAACAAATCCGCAATTCGCAATTAAGCAATGATGCTATATTAACGGCAAG 600
Db 541 GGCATGCAACAAATCCGCAATTCGCAATTAAGCAATGATGCTATATTAACGGCAAG 600
QY 601 CTTGAGTAAAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 601 CTTGAGTAAAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 AGCAGCTACGGGTGTATGAGCAAGAGCAGATGTAAGGCTTCCGATACAGGCTTAT 720
Db 661 AGCAGCTACGGGTGTATGAGCAAGAGCAGATGTAAGGCTTCCGATACAGGCTTAT 720
QY 721 ACAGGTGCAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 ACAGGTGCAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCAATGGAAGGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 GCAATGGAAGGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGTTAAGAAAGGCTCCATTAATAAGAAATGCGCTCAGGCGAATCTAGTCTTC 900
Db 841 TCCGTTAAGAAAGGCTCCATTAATAAGAAATGCGCTCAGGCGAATCTAGTCTTC 900
QY 901 CAATCTATCATGATGATGCGGTGGGAGCTTGAGAGCACTACCTTGAATCTGCAACCTTA 960
Db 901 CAATCTATCATGATGATGCGGTGGGAGCTTGAGAGCACTACCTTGAATCTGCAACCTTA 960
QY 961 TTCAAGCAAGCATACGCTGCTGTCAGAAATTCATACAACTCTGAGGAGCAAGCTG 1020
Db 961 TTCAAGCAAGCATACGCTGCTGTCAGAAATTCATACAACTCTGAGGAGCAAGCTG 1020
QY 1021 AATGGGCTTACACACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1021 AATGGGCTTACACACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 AGCATCTTTTGGCTGCGGGAATGAAGACCGAAGCGGGAACATCACTGATGACCAAG 1140
Db 1081 AGCATCTTTTGGCTGCGGGAATGAAGACCGAAGCGGGAACATCACTGATGACCAAG 1140
QY 1141 ACAGCTAAATATGCAATTAACGATGAGCTACGAAATCTCGCCCAAGCTTTGGGCT 1200
Db 1141 ACAGCTAAATATGCAATTAACGATGAGCTACGAAATCTCGCCCAAGCTTTGGGCT 1200
QY 1201 TATGGGCAATATCAACATGAGCAAGTCTCTTCAAGTGAAGCAAGATGATGATGAT 1260
Db 1201 TATGGGCAATATCAACATGAGCAAGTCTCTTCAAGTGAAGCAAGATGATGATGATGAT 1260

Db 841 TCGGATTTAGGAAACGGCTCCACTAATTAAGAAATGCGGCTGAGGGAACTGAGTTTC 900
QY 901 CAATCTATCATGATATGCGGTGGGGAAGCTTGAGAGACTGATCTGAAATCTGCAAACTTAA 960
Db 901 CAATCTATCATGATATGCGGTGGGGAAGCTTGAGAGACTGATCTGAAATCTGCAAACTTAA 960
QY 961 TTGAGCGAAGCATACGTGTGCTGGGCAAGAAATTCATCAAACTCTGGGGAGAGAGAGT 1020
Db 961 TTGAGCGAAGCATACGTGTGCTGGGCAAGAAATTCATCAAACTCTGGGGAGAGAGAGT 1020
QY 1021 AATGGGCTTACACAAGATTCGAGAAATGAGATGATGATGATGATGATGATGATGATGATG 1080
Db 1021 AATGGGCTTACACAAGATTCGAGAAATGAGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 ACGATCTTTTCCCTGCGGGAAATGAAGGACCGAAGCGCGGAAACATCAAGTGACCAAGC 1140
Db 1081 ACGATCTTTTCCCTGCGGGAAATGAAGGACCGAAGCGCGGAAACATCAAGTGACCAAGC 1140
QY 1141 ACGATCTTTTCCCTGCGGGAAATGAAGGACCGAAGCGCGGAAACATCAAGTGACCAAGC 1200
Db 1141 ACGATCTTTTCCCTGCGGGAAATGAAGGACCGAAGCGCGGAAACATCAAGTGACCAAGC 1200
QY 1201 TATGCGGACATATCAACATGATGAGCAAGTTCTTCAAGTGAACGCAAGAAATGAATGA 1260
Db 1201 TATGCGGACATATCAACATGATGAGCAAGTTCTTCAAGTGAACGCAAGAAATGAATGA 1260
QY 1261 CGGATCAAAACCGGATGATGAGCAAGTTCTTCAAGTGAACGCAAGAAATGAATGAATGA 1320
Db 1261 CGGATCAAAACCGGATGATGAGCAAGTTCTTCAAGTGAACGCAAGAAATGAATGAATGA 1320
QY 1321 GCAACCGGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1380
Db 1321 GCAACCGGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1380
QY 1321 GCAACCGGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1380
Db 1321 GCAACCGGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1380
QY 1381 TCCATGAGCTACACGATGCTGAGCAAGTTCTTCAAGTGAACGCAAGAAATGAATGAATGA 1440
Db 1381 TCCATGAGCTACACGATGCTGAGCAAGTTCTTCAAGTGAACGCAAGAAATGAATGAATGA 1440
QY 1441 AACAGAGGATCAACCAAGCTTCTTCAATTAAGGCGGACGATGTCGGGTGACGCT 1500
Db 1441 AACAGAGGATCAACCAAGCTTCTTCAATTAAGGCGGACGATGTCGGGTGACGCT 1500
QY 1501 GACATGCGCTTGGCTACCCGGAACGCTTCAAGTGAACGCAAGTGAACGCAAGTGAACGCA 1560
Db 1501 GACATGCGCTTGGCTACCCGGAACGCTTCAAGTGAACGCAAGTGAACGCAAGTGAACGCA 1560
QY 1561 TCCCTGAACGCTTACATGAGCAAGTCAAGTTCTTCAAGTGAACGCAAGTGAACGCAAGTGA 1620
Db 1561 TCCCTGAACGCTTACATGAGCAAGTCAAGTTCTTCAAGTGAACGCAAGTGAACGCAAGTGA 1620
QY 1621 TACTGCTTACTGCTACGCGGCAAGCTTGAAGAAATCTCCCTGGATAGCTGATGCTGATGCT 1680
Db 1621 TACTGCTTACTGCTACGCGGCAAGCTTGAAGAAATCTCCCTGGATAGCTGATGCTGATGCT 1680
QY 1681 CCTGCGAGCAACATGCTTCCGTAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1740
Db 1681 CCTGCGAGCAACATGCTTCCGTAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1740
QY 1741 CCAATATGAGCAACATGATGAGAAATGACTTACTTCTGCAATGATGATGATGATGATGATG 1800
Db 1741 CCAATATGAGCAACATGATGAGAAATGACTTACTTCTGCAATGATGATGATGATGATGATG 1800
QY 1801 GGGCGGCAATACCTAGAAATGATATTAATTAATGATGATGATGATGATGATGATGATGATG 1860
Db 1801 GGGCGGCAATACCTAGAAATGATATTAATTAATGATGATGATGATGATGATGATGATGATG 1860
QY 1861 GAGGTACAGGCTTATTAAGTACGCTTGAACGCAAGCTTCTGCTTGGCAATTTGGAAT 1920
Db 1861 GAGGTACAGGCTTATTAAGTACGCTTGAACGCAAGCTTCTGCTTGGCAATTTGGAAT 1920
QY 1921 TAA 1923
Db 1921 TAA 1923

Db 1921 TAA 1923
RESULT 6
US-09-920-954-7
Sequence 7, Application US/0920954
Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKI, KATSUSHI
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGIYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
PRIORITY FILING DATE: 2001-08-03
PRIORITY FILING DATE: 09/509,814
PRIORITY FILING DATE: 2000-04-06
PRIORITY FILING DATE: 2000-04-06
PRIORITY FILING DATE: 1998-10-07
PRIORITY FILING DATE: 1997-06-08
PRIORITY FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-09-920-954-7
Query Match 99.8%; Score 1919.8; DB 3; Length 1923;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ARGAGAAAGAAAGAAAGGTTTATCTGTTTATCGCTGAGCTGAGAGATTTGTGACT 60
Db 1 ARGAGAAAGAAAGAAAGGTTTATCTGTTTATCGCTGAGCTGAGAGATTTGTGACT 60
QY 61 GTTGCGTTAATATTCATCTGAGGTGTCAGAGAAATTTGATCTGATTTCAAGAA 120
Db 61 GTTGCGTTAATATTCATCTGAGGTGTCAGAGAAATTTGATCTGATTTCAAGAA 120
QY 121 ATTGAGACAACTGATGCTAAAGTTTCTCCAGAGCGGCGAAGCTGTGCTGCT 180
Db 121 ATTGAGACAACTGATGCTAAAGTTTCTCCAGAGCGGCGAAGCTGTGCTGCT 180
QY 181 TTTTGTGTGATGCTGAAATGTCGAAATCCCAAAAGTTTTCAGAGAAAGCTTGAAGA 240
Db 181 TTTTGTGTGATGCTGAAATGTCGAAATCCCAAAAGTTTTCAGAGAAAGCTTGAAGA 240
QY 241 GTTCCGCAATTAATTAATCTCAATTAATCAATTAATCAATTAATTAATTAATTAATTA 300
Db 241 GTTCCGCAATTAATTAATCTCAATTAATCAATTAATCAATTAATTAATTAATTAATTA 300
QY 301 AAAAGCAAGCTGAAAGAAAGAGGCGAAGATTTCTGCACTACATACCTGATTAATGCTTAC 360
Db 301 AAAAGCAAGCTGAAAGAAAGAGGCGAAGATTTCTGCACTACATACCTGATTAATGCTTAC 360
QY 361 ATTGTGATTAAGAGGCGAATGTTAAGTGAAGCAAGCAATTAAGAGAGCTGGAATTC 420
Db 361 ATTGTGATTAAGAGGCGAATGTTAAGTGAAGCAAGCAATTAAGAGAGCTGGAATTC 420
QY 421 GTGAGAGCTTATTTGCGGATTAATGAGATTAATGATCCCAAGCTTTTCAAGAAAGGCGATCA 480
Db 421 GTGAGAGCTTATTTGCGGATTAATGAGATTAATGATCCCAAGCTTTTCAAGAAAGGCGATCA 480

QY 481 GAGCTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAAAATAAGAGGTGCAATTAAAG 540
DB 481 GAGCTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAAAATAAGAGGTGCAATTAAAG 540
QY 541 GGCATGGAACAAATCGCACAATTCGCATTAAGCAATGATGTCTATATATTAACGCAAG 600
DB 541 GGCATGGAACAAATCGCACAATTCGCATTAAGCAATGATGTCTATATATTAACGCAAG 600
QY 601 CCTGAGATTAAGGTGATGAAATGATGTGGCGCGGGAATTTGCAACCGCATGTGCTCAG 660
DB 601 CCTGAGATTAAGGTGATGAAATGATGTGGCGCGGGAATTTGCAACCGCATGTGCTCAG 660
QY 661 AGCAGCTACGCGTGTGATGACAAAGCAGATGTAAGCGGTTCGCGATACAGGCGTTGAT 720
DB 661 AGCAGCTACGCGTGTGATGACAAAGCAGATGTAAGCGGTTCGCGATACAGGCGTTGAT 720
QY 721 ACAGGTGCAATGACAGTTGATGATGAAAGCCTTCGCGGGAATACTGCAATTATAT 780
DB 721 ACAGGTGCAATGACAGTTGATGATGAAAGCCTTCGCGGGAATACTGCAATTATAT 780
QY 781 GCATTGGGACGGAAGCAATTAAGCAATGATGCAATGATGCAATGATGCTGGC 840
DB 781 GCATTGGGACGGAAGCAATTAAGCAATGATGCAATGATGCAATGATGCTGGC 840
QY 841 TCGATTTAGAAACGCGCTCCAATAAAGAAATGCGCGCTCAGGCGAATCTAGCTTC 900
DB 841 TCGATTTAGAAACGCGCTCCAATAAAGAAATGCGCGCTCAGGCGAATCTAGCTTC 900
QY 901 CAATTTATCAATGATGCGGTGGGGAATTGAGAGCTACCTTCGAAATTCGAACTTTA 960
DB 901 CAATTTATCAATGATGCGGTGGGGAATTGAGAGCTACCTTCGAAATTCGAACTTTA 960
QY 961 TTCAGCCAGCATACAGTGTGGTGCAGAAATTCATACAACTCCGCGGAGCAGCAGTG 1020
DB 961 TTCAGCCAGCATACAGTGTGGTGCAGAAATTCATACAACTCCGCGGAGCAGCAGTG 1020
QY 1021 AATGGGGCTTACACACAGATTCACAGAAATGTGATGATATGTGCGCAAAATATATG 1080
DB 1021 AATGGGGCTTACACACAGATTCACAGAAATGTGATGATATGTGCGCAAAATATATG 1080
QY 1081 AGGATTCCTTTGGCTGCGCGGAAATGAAAGCAGAACCGCGGAAACCATCATGTCACAGGC 1140
DB 1081 AGGATTCCTTTGGCTGCGCGGAAATGAAAGCAGAACCGCGGAAACCATCATGTCACAGGC 1140
QY 1141 ACAGCTAAATATGCAATTAACAGTGGAGCTACGGAACCTCGCGCAAGCTTTGGGCT 1200
DB 1141 ACAGCTAAATATGCAATTAACAGTGGAGCTACGGAACCTCGCGCAAGCTTTGGGCT 1200
QY 1201 TATGCGGACATATCAACCATGTGCAAGTTCTCTTCACTGGAACGCAAAAGATGGA 1260
DB 1201 TATGCGGACATATCAACCATGTGCAAGTTCTCTTCACTGGAACGCAAAAGATGGA 1260
QY 1261 CGGATCAAAACCGGATGTCTATGCAACCGGAAAGCTTCACTATCAAGCAAGATCTTCTCT 1320
DB 1261 CGGATCAAAACCGGATGTCTATGCAACCGGAAAGCTTCACTATCAAGCAAGATCTTCTCT 1320
QY 1321 GCAACCGGATTCCTCTTCTGGGCGAACCATGCAATTAATATGCAATGAGGTGAAAG 1380
DB 1321 GCAACCGGATTCCTCTTCTGGGCGAACCATGCAATTAATATGCAATGAGGTGAAAG 1380
QY 1381 TCCATGCGTACACCCATCGTTGCTGGAACGTTGCAAGCTTCGTAAGCATTTTGTGAA 1440
DB 1381 TCCATGCGTACACCCATCGTTGCTGGAACGTTGCAAGCTTCGTAAGCATTTTGTGAA 1440
QY 1441 AACAGAGCATACACCCAAAGCTTCTCTATTAAGGCGCATGATGCGGTGCAAGT 1500
DB 1441 AACAGAGCATACACCCAAAGCTTCTCTATTAAGGCGCATGATGCGGTGCAAGT 1500
QY 1501 GACATGCGCTTGGCTACCCGAAACGGTAAACAAAGATGGGGAAGATGATGATTA 1560
DB 1501 GACATGCGCTTGGCTACCCGAAACGGTAAACAAAGATGGGGAAGATGATGATTA 1560
QY 1561 TCCCTGAACGTTGCTATGTGAACGATTCAGTTCTCTATTCACCAAGCCAAAGACG 1620

DB 1561 TCCCTGAACGTTGCTATGTGAACGATTCAGTTCTCTATTCACCAAGCCAAAGACG 1620
QY 1621 TACTGTTTACTGTACTACGCGGCAAGCCTTTGAAATCTCCCTGTAATGTCTGATGCC 1680
DB 1621 TACTGTTTACTGTACTACGCGGCAAGCCTTTGAAATCTCCCTGTAATGTCTGATGCC 1680
QY 1681 CCTGGACACAACTGCTTCCGTAACGCTTGTCAATGATCTGACCTTGTCAATACCGCT 1740
DB 1681 CCTGGACACAACTGCTTCCGTAACGCTTGTCAATGATCTGACCTTGTCAATACCGCT 1740
QY 1741 CCAATGCGACACAGTATGTAAGAAATGACCTTACCTTGCCCATCAATGATTAACCGGAT 1800
DB 1741 CCAATGCGACACAGTATGTAAGAAATGACCTTACCTTGCCCATCAATGATTAACCGGAT 1800
QY 1801 GCGCGCAATTAACGTAAGAAATGATTTATTAATGACCAACAAAGCGGACGTATACAT 1860
DB 1801 GCGCGCAATTAACGTAAGAAATGATTTATTAATGACCAACAAAGCGGACGTATACAT 1860
QY 1861 GAGGTACAGGCTTATTAACGTAACCGGTGGAACCAAGACCTTCTGTTGGCAATGTGAAT 1920
DB 1861 GAGGTACAGGCTTATTAACGTAACCGGTGGAACCAAGACCTTCTGTTGGCAATGTGAAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923

RESULT 7
US-10-784-870-7
; Sequence 7, Application US/10784870
; Publication No. US20040142837A1
; GENERAL INFORMATION:
; APPLICANT: TAKAWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/10/784, 870
; PRIOR FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 7
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-10-784-870-7

Query Match 99.8%; Score 1919.8; DB 7; Length 1923;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCACTGCAAGATTTTGTGCACT 60
DB 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCACTGCAAGATTTTGTGCACT 60
QY 61 GTTGGTTAAGTATCATCTTGCAGGTGTGCAAGAAATTTGATCTGATTTCAAGGA 120
DB 61 GTTGGTTAAGTATCATCTTGCAGGTGTGCAAGAAATTTGATCTGATTTCAAGGA 120

121 ATTGAGCAAACTGATGCTAAAGGTTTCTCAAGAGGGGGGAGAGCTGGTCTGCT 180
121 ATTGAGCAAACTGATGCTAAAGGTTTCTCAAGAGGGGGGAGAGCTGGTCTGCT 180
181 TTTCTGGTGAATCTGAAAAATGTGAAACTCCCAAAAGGTTTGAGAAAGGCTTGAAA 240
181 TTTCTGGTGAATCTGAAAAATGTGAAACTCCCAAAAGGTTTGAGAAAGGCTTGAAA 240
241 GTCCCGGCAAAATTAATTAATCTCATATTAATCAATTCATGAGCAATTTTGAAGAAA 300
241 GTCCCGGCAAAATTAATTAATCTCATATTAATCAATTCATGAGCAATTTTGAAGAAA 300
301 AAACAGAGCTGGAAAAAAGAGGGGCAAAAGTTCTGACATACATACCTGATTAAGCTTAC 360
301 AAACAGAGCTGGAAAAAAGAGGGGCAAAAGTTCTGACATACATACCTGATTAAGCTTAC 360
361 ATTGTCAGATATGAGGGGCGATGTTAAGTCAGCAACAAGCAACATTCAGCAAGTGAATCC 420
361 ATTGTCAGATATGAGGGGCGATGTTAAGTCAGCAACAAGCAACATTCAGCAAGTGAATCC 420
421 GTGAGGCTTATTTCCGATATACAGAAATGATCCCAAGCTTTTCAAAAAGGGGCAATCA 480
421 GTGAGGCTTATTTCCGATATACAGAAATGATCCCAAGCTTTTCAAAAAGGGGCAATCA 480
481 GAGCTGTAAAGAGTGGGGCTTGATACAAAGGAAAAATTAAGAGTGCATTTAAGA 540
481 GAGCTGTAAAGAGTGGGGCTTGATACAAAGGAAAAATTAAGAGTGCATTTAAGA 540
481 GAGCTGTAAAGAGTGGGGCTTGATACAAAGGAAAAATTAAGAGTGCATTTAAGA 540
541 GGCATCGAACAATTCGCAATTCGCAATTAAGCAATGATGCTATATTAATTAACGGCAAG 600
541 GGCATCGAACAATTCGCAATTCGCAATTAAGCAATGATGCTATATTAATTAACGGCAAG 600
541 GGCATCGAACAATTCGCAATTCGCAATTAAGCAATGATGCTATATTAATTAACGGCAAG 600
601 CTTGAGTAAAGTGAATGAATGATGTTGCGGTGGAATGTCAAAGGGGATGTCAG 660
601 CTTGAGTAAAGTGAATGAATGATGTTGCGGTGGAATGTCAAAGGGGATGTCAG 660
601 CTTGAGTAAAGTGAATGAATGATGTTGCGGTGGAATGTCAAAGGGGATGTCAG 660
661 AGCAGCTAAGGGGTTGATGATGCAAGAGCAGATGCTGAGGGTTTCCGATTAAGGGCTTGAT 720
661 AGCAGCTAAGGGGTTGATGATGCAAGAGCAGATGCTGAGGGTTTCCGATTAAGGGCTTGAT 720
661 AGCAGCTAAGGGGTTGATGATGCAAGAGCAGATGCTGAGGGTTTCCGATTAAGGGCTTGAT 720
721 ACAGGTCGCAATGACATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
721 ACAGGTCGCAATGACATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
721 ACAGGTCGCAATGACATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
781 GCAATGGGACGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
781 GCAATGGGACGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
841 TCCGATTAAGGAAACGGCTCCACTTAATTAAGGAAATGCGCTCAAGGGGATTAAGCTTTC 900
841 TCCGATTAAGGAAACGGCTCCACTTAATTAAGGAAATGCGCTCAAGGGGATTAAGCTTTC 900
901 CAATCATCATGATAGCGGTGGGAGCTTGAAGAGCTACCTTGAATCTGCAAACTTGA 960
901 CAATCATCATGATAGCGGTGGGAGCTTGAAGAGCTACCTTGAATCTGCAAACTTGA 960
901 CAATCATCATGATAGCGGTGGGAGCTTGAAGAGCTACCTTGAATCTGCAAACTTGA 960
961 TTGAGCCAAAGCATACAGTGTGTCGAGAAATTCATCAAACTCTTGAGGAGCAGCAGT 1020
961 TTGAGCCAAAGCATACAGTGTGTCGAGAAATTCATCAAACTCTTGAGGAGCAGCAGT 1020
961 TTGAGCCAAAGCATACAGTGTGTCGAGAAATTCATCAAACTCTTGAGGAGCAGCAGT 1020
1021 AATGGGGCTTACACACATTCGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1080
1021 AATGGGGCTTACACACATTCGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1080
1021 AATGGGGCTTACACACATTCGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1080
1081 ACAGTCTTTTCTGCTGCGGGAATGAGAGCCGAGCGGGAACATCAATGAGCAAGG 1140
1081 ACAGTCTTTTCTGCTGCGGGAATGAGAGCCGAGCGGGAACATCAATGAGCAAGG 1140
1081 ACAGTCTTTTCTGCTGCGGGAATGAGAGCCGAGCGGGAACATCAATGAGCAAGG 1140
1141 ACAGTCTTTTCTGCTGCGGGAATGAGAGCCGAGCGGGAACATCAATGAGCAAGG 1200
1141 ACAGTCTTTTCTGCTGCGGGAATGAGAGCCGAGCGGGAACATCAATGAGCAAGG 1200
1141 ACAGTCTTTTCTGCTGCGGGAATGAGAGCCGAGCGGGAACATCAATGAGCAAGG 1200

1201 TATGCGGCAATATATCAACATGATGAGCAAGTTCCTTCAAGTGAACCGAAGGATGGA 1260
1201 TATGCGGCAATATATCAACATGATGAGCAAGTTCCTTCAAGTGAACCGAAGGATGGA 1260
1261 CGGATCAAAACCGATGTCATGAGCAACGGGAAACCTTCACTATCAGCAAGATCTTCTT 1320
1261 CGGATCAAAACCGATGTCATGAGCAACGGGAAACCTTCACTATCAGCAAGATCTTCTT 1320
1321 GCAACCGGATTCCTCTTCTGAGGCAACATGACAGTAAATATGATCATGAGTGAACG 1380
1321 GCAACCGGATTCCTCTTCTGAGGCAACATGACAGTAAATATGATCATGAGTGAACG 1380
1321 GCAACCGGATTCCTCTTCTGAGGCAACATGACAGTAAATATGATCATGAGTGAACG 1380
1381 TCCATGCGTACACAGATTCGTTGCGGAAAGTGGGCAAGCTTGTGAGCAATTTTGA 1440
1381 TCCATGCGTACACAGATTCGTTGCGGAAAGTGGGCAAGCTTGTGAGCAATTTTGA 1440
1381 TCCATGCGTACACAGATTCGTTGCGGAAAGTGGGCAAGCTTGTGAGCAATTTTGA 1440
1441 AACAGAGGCAATCAACCAAGCTTCTATTAAGAGCGGCACTGATTCGGGTCAGCT 1500
1441 AACAGAGGCAATCAACCAAGCTTCTATTAAGAGCGGCACTGATTCGGGTCAGCT 1500
1441 AACAGAGGCAATCAACCAAGCTTCTATTAAGAGCGGCACTGATTCGGGTCAGCT 1500
1501 GACATGCGCTTGGCTACCCGAAACGGTAAACAGAGTGGGAGCAGTGAATGATTA 1560
1501 GACATGCGCTTGGCTACCCGAAACGGTAAACAGAGTGGGAGCAGTGAATGATTA 1560
1501 GACATGCGCTTGGCTACCCGAAACGGTAAACAGAGTGGGAGCAGTGAATGATTA 1560
1561 TCCCTGAACGTTGCTATGATGAGACAGTCCAGTTCCTATACCAAGCCCAAAAGCAG 1620
1561 TCCCTGAACGTTGCTATGATGAGACAGTCCAGTTCCTATACCAAGCCCAAAAGCAG 1620
1561 TCCCTGAACGTTGCTATGATGAGACAGTCCAGTTCCTATACCAAGCCCAAAAGCAG 1620
1621 TACTGTTTACTGCTACTGCGGCAAGCTTGAAGATCCCTGATATGATGATGATGATGAT 1680
1621 TACTGTTTACTGCTACTGCGGCAAGCTTGAAGATCCCTGATATGATGATGATGATGAT 1680
1621 TACTGTTTACTGCTACTGCGGCAAGCTTGAAGATCCCTGATATGATGATGATGATGAT 1680
1681 CTTGCGAGCAACCTGCTTCCGTAACGCTTGTCAATGATGAGCACTTGTCAATGAGCT 1740
1681 CTTGCGAGCAACCTGCTTCCGTAACGCTTGTCAATGATGAGCACTTGTCAATGAGCT 1740
1681 CTTGCGAGCAACCTGCTTCCGTAACGCTTGTCAATGATGAGCACTTGTCAATGAGCT 1740
1741 CCAATGGGCAACAGATGATGAGAAATGATCTTACTTCCGATTAAGGATTAATGATGAT 1800
1741 CCAATGGGCAACAGATGATGAGAAATGATCTTACTTCCGATTAAGGATTAATGATGAT 1800
1741 CCAATGGGCAACAGATGATGAGAAATGATCTTACTTCCGATTAAGGATTAATGATGAT 1800
1801 GGCAGCAATACGTAAGAAATGATTAATTAATGACCAACAAAGCGGAGCTTATCAAT 1860
1801 GGCAGCAATACGTAAGAAATGATTAATTAATTAATGACCAACAAAGCGGAGCTTATCAAT 1860
1801 GGCAGCAATACGTAAGAAATGATTAATTAATTAATGACCAACAAAGCGGAGCTTATCAAT 1860
1861 GAGGTAACGCTTATTAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1861 GAGGTAACGCTTATTAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1861 GAGGTAACGCTTATTAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1921 TAA 1923
1921 TAA 1923
1921 TAA 1923

RESULT 8
US-09-920-954-3
; Sequence 3, Application US/0920954
; Publication No. US20020064854A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIKA
; APPLICANT: KOBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528

/ PRIOR FILING DATE: 1998-10-07
/ PRIOR APPLICATION NUMBER: JP 9-274570
/ PRIOR FILING DATE: 1997-06-08
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 3
/ LENGTH: 1920
/ TYPE: DNA
/ ORGANISM: Bacillus sp.
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1920)
US-09-954-3

Query Match 76.8%; Score 1477.2; DB 3; Length 1920;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 161; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 10 AAGAAAGAGTGTGTTTATCTGTTTATCAGCTGCAGCGATTTTGTGACTGTGCGTTA 69
DB 7 AAGAAAGAGTGTGTTTATCTGTTTATCAGCTGCAGCGATTTTGTGACTGTGCGTTA 66
QY 70 AGTAATCCATCTGCAGCTGTGCGAAGATTTGATCTGATTTCAAAGAAATTCAGACA 129
DB 67 AACCAATCCCTCGCGCTGTGATGCAAGACTTTTGTGATCTGATTTTAAAGAAATTCAGACA 126
QY 130 ACAACTGATGCTAAAGTCTTCTCCAGCAGCGGCGACATGCTGCTGCTTTCTGCTG 189
DB 127 ACAACGATGCTAGTGTGTTTCTCCAGCAGCGAAGAGTGTGCTGCTGCTTTCTGCTG 186
QY 190 GAATCTGAAATGTGAACTCCCAAAAGTTTGCAGAGAGAGCTTGAAACAGTCCGCGCA 249
DB 187 GAGTCTGAAATGTGAACTTTCTTAAGAGTTGCTTAAGAACTTGAAACAGTCCGCGCA 246
QY 250 AATAATAACTGCTATTTATCCAAATTCATGCAACCAATTTTGAAGAAACAAACAGCAG 309
DB 247 AATAATAACTGCTATTTATCCAAATTCATGCAACCAATTTTGAAGAAACAAACAGCAG 306
QY 310 CTGGAAGAAAGAGGCGCAAGATTTCTGACTACATCACTGATTTATGCTTACATTTGCGAG 369
DB 307 CTAGAGACAACTGAGAGCAAGATTTCTGACTACATCTGATTTATGCTTATTTGCGAG 366
QY 370 TATGAGGCGGATGTTAAGTCAAGCAACCAACCATTTGAGCAGTGTGAATCCGTGAGCCT 429
DB 367 TATGAGGCGGATGTTCACTCAAAAGTCCGCTCATTTGAACACGTGGAATCACTGAGCGCA 426
QY 430 TATTTGCCGATATACGAATAGATCCCGAGCTTTTCAAAAGGCGCATCAAGCTTTGA 489
DB 427 TACTTGCCTGAATATCAAAATAGATCCCGAGCTTTTCAAAAGGCGCATCAAGCTTTG 486
QY 490 AAGCAGTGGCGCTGATCAAAAGCAGAAATTAAGAGTGCATTAAGAGCATCGAA 549
DB 487 AAGCAGTGGCGCTGATCAAAAGCAGAAATTAAGAGTGCATTAAGAGCATCGAA 546
QY 550 CAAATGCAACATTCGCAATTAAGCAATGATGCTATATATTAACGCCAAAGCCTGATAT 609
DB 547 GAAATGCTCAGATAGCTAGCAAGCAATGACGTCTATATATTAACGCCAAAGCCTGATAT 606
QY 610 AAGGATGATGAATGATGTTGCGGTGGAATTTGCAAAAGCGAGTGTGCTGAGAGCTTAC 669
DB 607 AAGGATGATGAATGATGTTGCGGTGGAATTTGCAAAAGCGAGTGTGCTGAGAGCTTAC 666
QY 670 GGGTGTATGACAAAGACAGATCGTAGCGGTGCGATACAGAGGCTTGAATACAGGTGCG 729
DB 667 GGGTGTATGACAAAGCGCGAGATTTGCGAGTTGCGATACAGAGTGTGATACAGAGGAG 726
QY 730 AATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
DB 727 AATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
QY 790 CGGAGGATTAATGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
DB 787 CGGAGGATTAATGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846

QY 850 GGAAGCGGCTCCTAATTAAGGAATGGCGCTCAGCGGAATTAAGTCTTCCATCTATC 909
DB 847 GGAAGCGGCTCCTAATTAAGGAATGGCGCTCAGCGGAATTAAGTCTTCCATCTATC 906
QY 910 ATGATAGCGGTGGGAGCTTGAAGACCTAATGAACTTGAACCTTATTCAGCCAA 969
DB 907 ATGATAGCGGTGGGAGCTTGAAGACCTAATGAACTTGAACCTTATTCAGCCAA 966
QY 970 GCATACAGTGTGTGCGCAATTTATCAAACTCTTGGGAGCAGAGTGAATGGGCT 1029
DB 967 GCATACAGTGTGTGCGCAATTTATCAAACTCTTGGGAGCAGAGTGAATGGGCT 1026
QY 1030 TACACACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1089
DB 1027 TACACACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1086
QY 1090 TTGCTGCGGGAATGAAGACCGAACCGGGAACCATCACTGACACAGCAGCTTAA 1149
DB 1087 TTGCTGCGGGAATGAAGACCGAACCGGGAACCATCACTGACACCTGTAACGGCTTAA 1146
QY 1150 AATGCAATTAAGTGTGAGCTACGGAACCTCCGCGCACTTTGGGTCTTAATGCGAG 1209
DB 1147 AATGCAATTAAGTGTGAGCTACGGAACCTCCGCGCACTTTGGGTCTTAATGCGAG 1206
QY 1210 AATATCAACGATGTGCAAGTCTCTTCACTGATGATGATGATGATGATGATGATGAT 1269
DB 1207 AATATTAACGATGTGCAAGTCTCTTCACTGATGATGATGATGATGATGATGATGAT 1266
QY 1270 CCGATGTGATGAGCAGCGGGAACCTTCACTATCACTGATGATGATGATGATGATGAT 1339
DB 1267 CCGATGTGATGAGCAGCGGGAACCTTCACTATCACTGATGATGATGATGATGATGAT 1326
QY 1330 TCTCTCTTGTGGGGAACATGACATTAATGATGATGATGATGATGATGATGATGATGAT 1389
DB 1327 TCTCTCTTGTGGGGAACATGACATTAATGATGATGATGATGATGATGATGATGATGAT 1386
QY 1390 ACACGATGTTGTGGAAGAGTGGCAGAGCTGTGAGCACTTTTGTGAAACAGAGGC 1449
DB 1387 ACACGATGTTGTGGAAGAGTGGCAGAGCTGTGAGCACTTTTGTGAAACAGAGGC 1446
QY 1450 ATCAACCAAGACCTTCTCTATTAAGCGGCACTGATGATGATGATGATGATGATGAT 1509
DB 1447 ATCACTCTTAAGCCTTCTCTATTAAGCGGCACTGATGATGATGATGATGATGATGAT 1506
QY 1510 CTGCGCTACCGGAACGATTAACAGATGAGGAGAGATGATGATGATGATGATGATGAT 1569
DB 1507 CTGCGCTACCGGAACGATTAACAGATGAGGAGAGATGATGATGATGATGATGATGAT 1566
QY 1570 GTTGCCTATGGAAGAGTCCAGTCTCTATCAACAGCCAAAGCAAGCACTGATGAT 1629
DB 1567 GTTGCCTATGGAAGAGTCCAGTCTCTATCAACAGCCAAAGCAAGCACTGATGAT 1626
QY 1630 ACTGCTATGCGGCAAGCCTTTGAAATCTCCCTGATGATGATGATGATGATGATGATGAT 1689
DB 1627 ACTGCTATGCGGCAAGCCTTTGAAATCTCCCTGATGATGATGATGATGATGATGATGAT 1686
QY 1690 ACACTGCTTCCGTAAGCCTTGTCAATGATGATGATGATGATGATGATGATGATGATGAT 1749
DB 1687 ACACTGCTTCCGTAAGCCTTGTCAATGATGATGATGATGATGATGATGATGATGATGAT 1746
QY 1750 ACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1809
DB 1747 ACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1806
QY 1810 AAGGTAGAAATGATATTAATTAATGACCAACAAAGCGGAGCTATTAATTAATTAATGAT 1869
DB 1807 AAGGTAGAAATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGAT 1866
QY 1870 GCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1923
DB 1867 GCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920

RESULT 9
US-10-784-870-3
Sequence 3, Application US/10784870
Publication No. US20040142837A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIKA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784,870
PRIOR FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 1920
TYPE: DNA
ORGANISM: *Bacillus* sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
US-10-784-870-3

Query Match 76.8%; Score 1477.2; DB 7; Length 1920;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 10 AAAAAAAAAAGTGTATCTGTTTATCAGTCGACGCGATTTTGTGACGTCGTGCTTGA 69
DB 7 AAGAGAAAGGTGTATCTGTTTATCAGTCGACGCGATTTTGTGACGTCGTGCTTGA 66
QY 70 AGTATTCATCTGCGAGTGTGCAAGGATTTTGTATCTGATTTTCAAAGGATTCAGACA 129
DB 67 AACCAATCCCTGCGGTGTATGCAAGGATTTTGTATCTGATTTTAAAGGATTCAGACA 126
QY 130 ACAACGTATGCTAAAGTGTCTCAAGCAGGGGCAAGCTGTGTGCTGCTTTCTGTG 189
DB 127 ACAACGTATGCTGTGTGTCTCAAGCAGGGGCAAGGCTGTGTGCTGCTTTCTGTG 186
QY 190 GAATGTGAAGATGTGAACCTCCCAAGGTTTGCAGAGAGAGCTTGAACAGTCCCGCA 249
DB 187 GAGTGTGAAGATGTGAACCTTAAAGGATTTGTGAAGAGAGCTTGAACAGTCCCGCA 246
QY 250 AATATTAAGTCTCATTTATCTCAATTCAGTGAAGCAATTTTGAAGAGAGCAAGCAG 309
DB 247 AATATTAAGTCTCATTTATCTCAATTCAGTGAAGCAATTTTGAAGAGAGCAAGCAG 306
QY 310 CTGAAAAAAGCAGGGGCAAGGATTCGACGCTACATACCGATTATGCTTACATTTGCGAG 369
DB 307 CTGAGACGACCTGAGAGCAAGGATTCGACGCTACATACCGATTATGCTTACATTTGCGAG 366
QY 370 TATGAGGGGAGTGTAAAGTCAAGCAAGACCATTTGAGCAGTGAATCCGTGAGCCT 429
DB 367 TATGAGGGGAGTGTCAAGTCAAGGATCCGCTCATTTGAACAGTGAATCACTGAGCCA 426
QY 430 TATTGCGGATTAAGGATTAAGTCCAGCTTTTCAAGAGGGGAGTCAAGGCTTGTGA 489
DB 427 TACTTGCAGAAATACAAATATGATCCCGCTTTTCAAGAGGGGAGTCAAGGCTTGTGA 486
QY 490 AAGGAGTGGGCTGTGATCAAGAGCAAGGATTAAGAGGTCATTTAAGAGGATCGAA 549

DB 487 AAGGCTGGCGCTGTATGACAGAGCAAGCAATTAAGAGTCATTAAGAGCATCGAG 546
QY 550 CAAATGCACAATTCGCAATTAAGCAATGATGTCTATATATTAATTCGCAAGGCTGAGTAT 609
DB 547 GAAATGCTCACTAGTACGAGCAAGCAATGATGTCTATATATTAATTCGCAAGGCTGAGTAT 606
QY 610 AAGGAGTAAATGATGTGGCGGTGAATTTGTCAAGCGGAGTGTCTCAAGGAGCTTAC 669
DB 607 AAGGAGTAAATGATGTGGCGGTGAATTTGTCAAGCGGAGTGTCTCAAGGAGCTTAC 666
QY 670 GGGTGTATGACAGAGCAGATCTGAGCGGTGCGGATCAAGGCTGTATCAAGTCCG 729
DB 667 GGTGTGTATGACAGAGCAGATTTGTGAGTTCGAGTTCGAGTATGATTAAGAGAA 726
QY 730 AATGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
DB 727 AACGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
QY 790 CCGAGGATTAAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
DB 787 CCGAGGATTAAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
QY 850 GGAAGCGGCTCACTAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 909
DB 847 GGAAGCGGCTCACTAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 906
QY 910 ATGATTAAGGCTGAGGAGCTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 969
DB 907 ATGATTAAGGCTGAGGAGCTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
QY 970 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
DB 967 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
QY 1030 TACACAGAGATTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
DB 1027 TACACAGAGATTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
QY 1090 TTGCTGCGGAGATGAAGAGCCGAGCGGAGCATCACTGATGACAGGCAAGCTTAA 1149
DB 1087 TTGCTGCGGAGATGAAGAGCCGAGCGGAGCATCACTGATGACAGGCAAGCTTAA 1146
QY 1150 AATGCAATTAAGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1209
DB 1147 AATGCAATTAAGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1206
QY 1210 AATATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1269
DB 1207 AATATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
QY 1270 CCGATGTATGAGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1329
DB 1267 CCGATGTATGAGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1326
QY 1330 TCTCTCTTGTGGGAGCAAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1389
DB 1327 TCTCTCTTGTGGGAGCAAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386
QY 1390 AACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1449
DB 1387 AACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
QY 1450 ATCAAGCAAGGCTTCTCTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1509
DB 1447 ATCAAGCAAGGCTTCTCTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1506
QY 1510 CTGGCTACCGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1569
DB 1507 CTGGCTACCGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1566
QY 1570 GTTGCCTATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1629
DB 1567 GTTGCCTATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626

QY 1630 ACTGCTACTGCGCGAAGCCTTTGAAATCTCCCTGTAATGCTGATGCCCTTCGAGC 1689
 DB 1627 ACTGCAACGGCGGCGAAGCATTTGAAATCTCCCTGTAATGCTGATGCCCTTCGAGC 1686
 QY 1690 ACAATGCTTCCTGTAACGCTTGTCAATGATCTGACCTTGTCAATACCGCTCCAAATGAC 1749
 DB 1687 ACTATGCTTCCTGTAACGCTTGTCAATGATCTGACCTTGTCAATACCGCTCCAAATGAC 1746
 QY 1750 ACACGTAATGTAAGAAATGATCTTACTTCCCAATGATGATGATGATGATGATGATGAT 1809
 DB 1747 ACMAATATGCTGGAATGATCTTCAAGACCAATTTGACATGATGATGATGATGATGAT 1806
 QY 1810 AACGTAAATGATATTTATTAATGACCAACGCGGACGTATACATTAATGATGATGATGAT 1869
 DB 1807 AACGTAAATGATATTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1866
 QY 1870 GCTTAAACGTACCGGTGACCAACGCTTCTGTTGGCAATTTGATTAATTA 1923
 DB 1867 GCATATATGTCGCGGTGACCAACGCTTCTGTTGGCAATTTGATTAATTA 1920

RESULT 10

US-10-385-662-1
 / Sequence 1, Application US/10385662
 / Publication No. US20040002432A1
 / GENERAL INFORMATION:
 / APPLICANT: OKUDA, MITSUYOSHI
 / APPLICANT: SATO, TSUYOSHI
 / APPLICANT: SAITO, KAZUHIRO
 / APPLICANT: SUMITOMO, NOBUYUKI
 / APPLICANT: IZAWA, YOSHIFUMI
 / APPLICANT: SAKKI, KATSUHIISA
 / APPLICANT: KOBAYASHI, TOHRU
 / APPLICANT: NOMURA, MASAFUMI
 / TITLE OF INVENTION: Alkaline protease
 / FILE REFERENCE: 234938US0
 / CURRENT APPLICATION NUMBER: US/10/385,662
 / CURRENT FILING DATE: 2003-03-12
 / PRIOR APPLICATION NUMBER: JP 2002-081428
 / PRIOR FILING DATE: 2002-03-22
 / PRIOR APPLICATION NUMBER: JP 2002-165987
 / PRIOR FILING DATE: 2002-06-06
 / PRIOR APPLICATION NUMBER: JP 2002-304230
 / PRIOR FILING DATE: 2002-10-18
 / PRIOR APPLICATION NUMBER: JP 2002-304231
 / PRIOR FILING DATE: 2002-10-18
 / NUMBER OF SEQ ID NOS: 20
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO 1
 / LENGTH: 1305
 / TYPE: DNA
 / ORGANISM: Bacillus sp. KSM-KP43
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (1)..(1305)
 / OTHER INFORMATION:
 / US-10-385-662-1

Query Match 67.9%; Score 1305; DB 6; Length 1305;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 619 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAAGACGACTACGGGTTGAT 678
 DB 1 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAAGACGACTACGGGTTGAT 60
 QY 679 GGAACAAGACAGATCTATCGGTTGCGGATACAGGGCTTGATACAGGTGCGAATGACAT 738
 DB 61 GGAACAAGACAGATCTATCGGTTGCGGATACAGGGCTTGATACAGGTGCGAATGACAT 120
 QY 739 TCGATGATGAAGCCTTCGCGGGAATTAATGATGATTAATGATGATGATGATGATGAT 798
 DB |||||

DB 121 TCGATGATGAAGCCTTCGCGGGAATTAATGATGATTAATGATGATGATGATGATGAT 180
 QY 799 AATGCAATGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 858
 DB 181 AATGCAATGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 859 TCCACTAATTAAGGAATGCGGCTCAAGCGGATGATGATGATGATGATGATGATGATGAT 918
 DB 241 TCCACTAATTAAGGAATGCGGCTCAAGCGGATGATGATGATGATGATGATGATGATGAT 300
 QY 919 GGTGGGGGACTTGGAGGACTACCTTCAATGCAATCTTATTCAGCGAAGCATAGAT 978
 DB 301 GGTGGGGGACTTGGAGGACTACCTTCAATGCAATCTTATTCAGCGAAGCATAGAT 360
 QY 979 GCTGGTGCAGAAATTCATACAACTCTGGGAGACAGATGAATGGGGCTTACACAA 1038
 DB 361 GCTGGTGCAGAAATTCATACAACTCTGGGAGACAGATGAATGGGGCTTACACAA 420
 QY 1039 GATTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
 DB 421 GATTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 1099 GGAATGAAAGACCGAACGGCGGAACCATCAATGACACAGGCAAGCTAAATGCAATA 1158
 DB 481 GGAATGAAAGACCGAACGGCGGAACCATCAATGACACAGGCAAGCTAAATGCAATA 540
 QY 1159 ACAGTGGAGCTACGGAACCTCTCGCCCAAGCTTTGGGTTTATGGGACAAATATCAAC 1218
 DB 541 ACAGTGGAGCTACGGAACCTCTCGCCCAAGCTTTGGGTTTATGGGACAAATATCAAC 600
 QY 1219 CATGTGGACAGTTCTCTTCAGTGAACCGAACCAAGATGAGACGGATCAACCGGATGTC 1278
 DB 601 CATGTGGACAGTTCTCTTCAGTGAACCGAACCAAGATGAGATCTTCTTGGACCGGATTCCTCTTC 720
 QY 1279 ATGGACCGGGAAGCTTCAATCAATCAAGAGATCTTCTTGGACCGGATTCCTCTTC 1338
 DB 661 ATGGACCGGGAAGCTTCAATCAATCAAGAGATCTTCTTGGACCGGATTCCTCTTC 720
 QY 1339 TGGGGAACCTGACAGTAATATGATATGATGATGATGATGATGATGATGATGATGATGAT 1398
 DB 721 TGGGGAACCTGACAGTAATATGATATGATGATGATGATGATGATGATGATGATGATGAT 780
 QY 1399 GTTGCTGGAACGATGACAGCTTCGTGAGATTTTGGAAAAACAGAGCATCAACCA 1458
 DB 781 GTTGCTGGAACGATGACAGCTTCGTGAGATTTTGGAAAAACAGAGCATCAACCA 840
 QY 1459 AAGCTTCTTATTAAGCGGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1518
 DB 841 AAGCTTCTTATTAAGCGGACATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 QY 1519 CCGAACGTAACCAAGATGAGGACGATGATGATGATGATGATGATGATGATGATGATGAT 1578
 DB 901 CCGAACGTAACCAAGATGAGGACGATGATGATGATGATGATGATGATGATGATGATGAT 960
 QY 1579 GTGAACGATCAGTTCTCTATCCACAGCCAAAAAGGAGATGATGATGATGATGATGAT 1638
 DB 961 GTGAACGATCAGTTCTCTATCCACAGCCAAAAAGGAGATGATGATGATGATGATGAT 1020
 QY 1639 GCCGGCAAGCTTTGAAAAATCTCCCTGATGATGATGATGATGATGATGATGATGATGAT 1698
 DB 1021 GCCGGCAAGCTTTGAAAAATCTCCCTGATGATGATGATGATGATGATGATGATGATGAT 1080
 QY 1699 TCCGTAAGCTTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758
 DB 1081 TCCGTAAGCTTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 QY 1759 GTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1818
 DB 1141 GTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 1819 AATGATTTATTAATGACCAAAAGCGGACGTATACATTAATGATGATGATGATGATGAT 1878
 DB 1201 AATGATTTATTAATGACCAAAAGCGGACGTATACATTAATGATGATGATGATGATGAT 1260

QY 1879 GTACCGGTTGGACCAAGACCTTCTCGTTGGCAATTGTAATTAA 1923
DB 1261 GTACCGGTTGGACCAAGACCTTCTCGTTGGCAATTGTAATTAA 1305

RESULT 11
US-10-820-712A-13
; Sequence 13, Application US/10820712A
; Publication No. US20050026804A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuoshi
; APPLICANT: Izawa, Yoshihumi
; APPLICANT: Kobayashi, Tohru
; APPLICANT: Koyama, Shingo
; APPLICANT: Sato, Tsuyoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 251701-USO
; CURRENT APPLICATION NUMBER: US/10/820,712A
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106708
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent version 3.2
; SEQ ID NO 13
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-9865
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: Protease 9865
US-10-820-712A-13

Query Match 66.7%; Score 1281.8; DB 8; Length 1302;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1289; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 619 AATGATGTTGGCGGTGGAATTTGCAAGCGGATGTGGCTCAGAGCAGTACGGGTTGAT 678
DB 1 AATGATGTTGGCGGTGGAATTTGCAAGCGGATGTGGCTCAGAGCAGTACGGGTTGAT 60
QY 679 GGAACAAGACAGATCGTACGGTGGCCGATACAGGGCTTGATACAGGTGGCAATGACAT 738
DB 61 GGAACAAGACAGATCGTACGGTGGCCGATACAGGGCTTGATACAGGTGGCAATGACAT 120
QY 739 TCGATGATGAGACCTTCCGCGGGAATAATTACATGATTAATGATGATGGAACGAGAT 798
DB 121 TCGATGATGAGACCTTCCGCGGGAATAATTACATGATTAATGATGGAACGAGAT 180
QY 799 AATGCCAATGATACGAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 858
DB 181 AATGCCAATGATACGAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 859 TCCATTAATTAAGAAATGCGCTTCAAGGGAATCTAGTTTTCGATCTATCATGATACG 918
DB 241 TCCATTAATTAAGAAATGCGCTTCAAGGGAATCTAGTTTTCGATCTATCATGATACG 300
QY 919 GGTGGGGGACTTGAAGACCTTCTGATCTGCAAACTTATTCGACCAAGATACATG 978
DB 301 GGTGGGGGACTTGAAGACCTTCTGATCTGCAAACTTATTCGACCAAGATACATG 360
QY 979 GGTGGTGGCAGAAATCATCAAACTCTGTTGGGAGCAGCAGTGAATGGGCTTACACACA 1038
DB 361 GGTGGTGGCAGAAATCATCAAACTCTGTTGGGAGCAGCAGTGAATGGGCTTACACACA 420
QY 1039 GATTTCAGAAATGTGATGATCTATGTGGCCAAATGATATGACGATCTTTTGGTGGC 1098
DB 421 GATTTCAGAAATGTGATGATCTATGTGGCCAAATGATATGACGATCTTTTGGTGGC 480
QY 1099 GGAATGAAGACCGAATCGGGAACCATCAGTGCACGAGCAAGCTAAATGCAATA 1158

DB 481 GGAATGAAGACCGAATCGGGAACCATCAGTGCACGAGCAAGCTAAATGCAATA 540
QY 1159 ACGTGGAGCTACGGAATACTCCGCCAAGCTTGGGTCTTATCGGACATATCAAC 1218
DB 541 ACGTGGAGCTACGGAATACTCCGCCAAGCTTGGGTCTTATCGGACATATCAAC 600
QY 1219 CATTGGGACAGTTCTTCACTGAGGACCGAAGAAAGATGACGAGTCAACCGGATGTC 1278
DB 601 CATTGGGACAGTTCTTCACTGAGGACCGAAGAAAGATGACGAGTCAACCGGATGTC 660
QY 1279 ATGCAACCGGGAACGTTCACTATCAGAGATCTTCTTGGACCGGATCTCTCTTC 1338
DB 661 ATGCAACCGGGAACGTTCACTATCAGAGATCTTCTTGGACCGGATCTCTCTTC 720
QY 1339 TGGGCGAACCATGACAGTAATATGATATCATGAGGTGAAAGCTCCATGCTACACGATC 1398
DB 721 TGGGCGAACCATGACAGTAATATGATATCATGAGGTGAAAGCTCCATGCTACACGATC 780
QY 1399 GTTGTGGAAGATGAGCAGCTTGTGAGCATTTTGTGAAAAAGAGGACATACACA 1458
DB 781 GTTGTGGAAGATGAGCAGCTTGTGAGCATTTTGTGAAAAAGAGGACATACACA 840
QY 1459 AAGCTTCTCTATTAATAAGCGGACATGATGCGGTGAGTGAATCGGCTTGGCTAC 1518
DB 841 AAGCTTCTCTATTAATAAGCGGACATGATGCGGTGAGTGAATCGGCTTGGCTAC 900
QY 1519 CCGAACGTTACCAAGATGGGGAAGATGACATTTGATTAATCCCTGAACGTTGCTAT 1578
DB 901 CCGAACGTTACCAAGATGGGGAAGATGACATTTGATTAATCCCTGAACGTTGCTAT 960
QY 1579 GTGAACGATCGAGTTCTTATCCACCAAGCCAAAAAGCAAGTATCTGTTACTACT 1638
DB 961 GTGAACGATCGAGTTCTTATCCACCAAGCCAAAAAGCAAGTATCTGTTACTACT 1020
QY 1639 GCCGCAAGCTTTGAAAAATCTCCGTGATGATGATGATGATGATGATGATGATGAT 1698
DB 1021 GCCGCAAGCTTTGAAAAATCTCCGTGATGATGATGATGATGATGATGATGATGAT 1080
QY 1699 TCCGTAACGTTGTCAATGATGTGACCTTGTCAATTAACGCTTCAATGACACAGTAT 1758
DB 1081 TCCGTAACGTTGTCAATGATGTGACCTTGTCAATTAACGCTTCAATGACACAGTAT 1140
QY 1759 GTGAAGATGATCTTATCTTCCGATACATGATTAATCTGGAATGCGGCAATAAGTAA 1818
DB 1141 GTGAAGATGATCTTATCTTCCGATACATGATTAATCTGGAATGCGGCAATAAGTAA 1200
QY 1819 AATGATTAATTAATGACACCAAGCGGGAAGTAACTTAATGAGGTACAGGCTTATAC 1878
DB 1201 AATGATTAATTAATGACACCAAGCGGGAAGTAACTTAATGAGGTACAGGCTTATAC 1260
QY 1879 GTACCGGTTGACCAAGACCTTCTGTTGGCAATTGTAATTAA 1919
DB 1261 GTACCGGTTGACCAAGACCTTCTGTTGGCAATTGTAATTAA 1301

RESULT 12
US-10-820-714A-14
; Sequence 14, Application US/10820714A
; Publication No. US20050214922A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuoshi
; APPLICANT: Kobayashi, Tohru
; APPLICANT: Sumitomo, Nobuyuki
; APPLICANT: Takimura, Yasuhiro
; APPLICANT: Sato, Tsuyoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 251697USO
; CURRENT APPLICATION NUMBER: US/10/820,714A
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106709
; NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-9865
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: Protease 9865
US-10-820-712A-14

Query Match 66.7%; Score 1281.8; DB 9; Length 1302;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1289; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
QY 619 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAGAGCAGTACGGGTGTAT 678
DB 1 AATGATGTTGCAAGTGTGCAAGCGGATGTGCGCAGAGCAGTACGGGTGTAT 60
QY 679 GGAAGAAGACAGATCGTAGCGGTGGCGATACAGGGCTTGATACAGGTGCGAATGACAGT 738
DB 61 GGAAGAAGACAGATCGTAGCGGTGGCGATACAGGGCTTGATACAGGTGCGAATGACAGT 120
QY 739 TCGATGATGAAGCCTTCCGCGGAAATTAATGATATGATGATGGAAGGAGAT 798
DB 121 TCGATGATGAAGCCTTCCGCGGAAATTAATGATATGATGGAAGGAGAT 180
QY 799 AATGCCAATGATACGAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 858
DB 181 AATGCCAATGATACGAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 859 TCCACTAATTAAGGAATGCGGCTCAGGGGAATCTAGTCTTCATCTCATATGATGATGAT 918
DB 241 TCCACTAATTAAGGAATGCGGCTCAGGGGAATCTAGTCTTCATCTCATATGATGATGAT 300
QY 919 GGTGGGGAAGCTTGAAGACTACCTTGAATCTGCAACCTTTATGCGCAAGCATACAGT 978
DB 301 GGTGGGGAAGCTTGAAGACTACCTTGAATCTGCAACCTTTATGCGCAAGCATACAGT 360
QY 979 GGTGGTGCAGAAATTCATACAACTCTGCGGAGACGACGATGAATGGGCTTACACACA 1038
DB 361 GGTGGTGCAGAAATTCATACAACTCTGCGGAGACGACGATGAATGGGCTTACACACA 420
QY 1039 GATTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
DB 421 GATTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 1099 GGGAAATGAAGACCGAAGCGGGAACATCACTGATGACACGACGATGATGATGATGAT 1158
DB 481 GGGAAATGAAGACCGAAGCGGGAACATCACTGATGACACGACGATGATGATGATGAT 540
QY 1159 ACAATGCGAAGCTTACGGAAGAACTCCGCGCAAGCTTGGGTCTTATGCGGACATATCAAC 1218
DB 541 ACAATGCGAAGCTTACGGAAGAACTCCGCGCAAGCTTGGGTCTTATGCGGACATATCAAC 600
QY 1219 CATGTGGAACAGTTCTCTTCAAGTGAACGACCAAGATGATGATGATGATGATGATGAT 1278
DB 601 CATGTGGAACAGTTCTCTTCAAGTGAACGACCAAGATGATGATGATGATGATGATGAT 660
QY 1279 AATGCAACCGGAAGCTTCAATCTATCAGCAAGATCTTCTTTCGACCGGATTCCTCTTC 1338
DB 661 AATGCAACCGGAAGCTTCAATCTATCAGCAAGATCTTCTTTCGACCGGATTCCTCTTC 720
QY 1339 TGGGGAACCAATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1398
DB 721 TGGGGAACCAATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 1399 GTTGTGGAACCGTGAACAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1458
DB 781 GTTGTGGAACCGTGAACAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 1459 AAGCTTCTTATTAAGGCGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1518
DB 1518 AAGCTTCTTATTAAGGCGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1518
```

```
DB 841 AAGCTTCTTATTAAGGCGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 1519 CCGAAGGTGAACCAAGATGAGGAGACAGTGAATGATGATGATGATGATGATGATGATGAT 1578
DB 901 CCGAAGGTGAACCAAGATGAGGAGACAGTGAATGATGATGATGATGATGATGATGATGAT 960
QY 1579 GTGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1638
DB 961 GTGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1639 GCCGGAAGCCTTGAAGAAATCTCCCTGATGATGATGATGATGATGATGATGATGATGAT 1698
DB 1021 GCCGGAAGCCTTGAAGAAATCTCCCTGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1699 TCCGTAAGCTTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758
DB 1081 TCCGTAAGCTTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1759 GTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1818
DB 1141 GTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1819 AATGATTTAATTAAGACCAAGGAGGAGCGATACATGATGATGATGATGATGATGAT 1878
DB 1201 AATGATTTAATTAAGACCAAGGAGGAGCGATACATGATGATGATGATGATGATGAT 1260
QY 1879 GTACCGGTTGACCAAGACCTTCTCGTTGCAATGATGATGATGATGATGATGATGAT 1919
DB 1261 GTACCGGTTGACCAAGACCTTCTCGTTGCAATGATGATGATGATGATGATGATGAT 1301

RESULT 13
US-10-820-712A-11
; Sequence 11, Application US/10820712A
; Publication No. US20050026804A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okada, Mitsuyoshi
; APPLICANT: Izawa, Yoshifumi
; APPLICANT: Kobayashi, Tohru
; APPLICANT: Koyama, Shingo
; APPLICANT: Sato, Tsuyoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; PILE REFERENCE: 251701-USO
; CURRENT APPLICATION NUMBER: US/10/820, 712A
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106708
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-kp9860
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: protease KP9860
US-10-820-712A-11

Query Match 51.8%; Score 997; DB 8; Length 1302;
Best Local Similarity 85.4%; Pred. No. 5.1e-287;
Matches 1111; Conservative 0; Mismatches 190; Indels 0; Gaps 0;
```

```
QY 619 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAGAGCAGTACGGGTGTAT 678
DB 1 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAGAGCAGTACGGGTGTAT 60
QY 679 GGAAGAAGACAGATCGTAGCGGTGGCGATACAGGGCTTGATACAGGTGCGAATGACAGT 738
DB 61 GGAAGAAGACAGATCGTAGCGGTGGCGATACAGGGCTTGATACAGGTGCGAATGACAGT 120
QY 739 TCGATGATGAAGCCTTCCGCGGAAATTAATGATATGATGATGATGATGATGATGATGATGAT 798
```

Db 121 TCGATGCAATGAAGCCTTCGGCGGTAAATAACAGCACTATATAGCACTGGGTGGAGCAAT 180
Qy 799 AATGCCAATGATACGAATGTGATGATGACCATGTGGCTCCGTATAGGAACGGC 858
Db 181 AATGGAAATGATACGAACGGTATGATCCCATGTGGCAAGTTCGGTATAGGAATGGC 240
Qy 859 TCCACTAATAAAGAAATGGCGCTCAGGGGAATCTAGTCTTCATCTATCTATGATAGC 918
Db 241 GCATGCAATTAAGAAATGGCACTCAGAGCAATCTGGTTCATCTATCTATGATAGC 300
Qy 919 GGTGGGGGACTTGAAGACTACCTTCGAATCTGCAACCTTATTCAGCAAGCATACAGT 978
Db 301 AGTGGGGGCTTGAAGGCTTGGCTTCATCTGCAACCTTATTCAGCAAGCATACAGT 360
Qy 979 GCTGTGCGCAATTCATCAAACTCTGGGGAGCAGCATGTAATGGGCTTACACACA 1038
Db 361 GCAGGTGCGCAATTCATCAAACTCTGGGGAGCAGCGGTGTAATGGGCTTACACACA 420
Qy 1039 GATTCCAGAAATGTGATGATCTATGTGCGCAAAATGATGACATCTTTCCTGCTGC 1098
Db 421 GATTCCAGAAATGTGATGATCTATGTGAGAAATGATGATGACATCTTTCCTGCTGC 480
Qy 1099 GGGAAATGAAGACCGAACCGGGAACCATCATGTCACAGGACAGCTTAAATGCAATA 1158
Db 481 GGGAAATGAAGACCGAACCGGGAACCATCATGTCACAGCTGTAACGGCTTAAACGGC 540
Qy 1159 ACAGTGCGAGTACGGAACCTTCGCGCCAGCTTGGGTCTTATGCGGACATATCAAC 1218
Db 541 ACTGTGCGGCGCAACCGAAACCTGCGGTCAAGCTTGGTTCCTATGCAATATATTAAC 600
Qy 1219 CATGTGCGCAATTCCTTCATGTCGACCGAAGAGATGACAGATCAAAACCGATGTC 1278
Db 601 CAGTGTGCAAGATTCCTTCCTGCGGCGCGAAGAAAGATGAGGCGAATCAAGCTGATGTC 660
Qy 1279 ATGGACCGGGAAAGCTTCTATCTATCATGAGCAAGATCTTCTTGGACCGGATCTCTTC 1338
Db 661 ATGGCGCCGAGGACATACATTTTATGAGCAAGATCTTCTTGGACCGGATCTCTTC 720
Qy 1339 TGGGCGAACCATGACATGATATATGATATGCGGTGAAACGTCCATGCTCAACCGATC 1398
Db 721 TGGGCGAATCATGACAGCAATATGCTATATGGGTGAAACGTCCATGCTCAACCGATC 780
Qy 1399 GTTGTGGAACATGTCACAGCTTGTGAGCATTTTGTGAAAAACAGAGCATCACACA 1458
Db 781 GTTGTGGAACATGTCACAGCTTGTGAGCATTTTGTGAAAAACAGAGCATCACACTCT 840
Qy 1459 AAGGCTTCTATTAATAAGCGGCACTGATTCGCGGTGAGCTGACATCGGCGCTGCTAC 1518
Db 841 AAGCTTCTCTATTTGAAGCAAGCTTGTGAGCTGCTGATGATGATGATGATGAT 900
Qy 1519 CGGAACGGTAAACGAAGATGGGAGCAAGTGAATGATGATGATGATGATGATGAT 1578
Db 901 CGGAACGGTAAACGAAGATGGGAGCAAGTGAATGATGATGATGATGATGATGAT 960
Qy 1579 GTGAACGAGTCAAGTCTCTATTCACAGCCAAAAACGAGCTACTGTTTACTGCTACT 1638
Db 961 GTGAACGAGTCAAGTCTCTATTCACAGCCAAAAACGAGCTACTGTTTACTGCAAG 1020
Qy 1639 GCGGCGAAGCCTTGAATAATCTCCGTGATGATGATGATGATGATGATGATGAT 1698
Db 1021 GCGGCGAAGCCTTGAATAATCTCCGTGATGATGATGATGATGATGATGATGAT 1080
Qy 1699 TCCGTAAAGCTTGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1758
Db 1081 TCTGTAAACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Qy 1759 GTAGAAATGATCTTATCTTGCATACAAATGATGATGATGATGATGATGATGAT 1818
Db 1141 GTGCGAAATGATCTTCTACAGCAATTTGCAATTAATGATGATGATGATGATGAT 1200
Qy 1819 AATGATTAATTAATGACCAAAAGCGGAGTATGATGATGATGATGATGATGAT 1878

Db 1201 AATGATTAATTAATTCGCCCCCAAGTGAACATATACATGAGTGAACATATAT 1260
Qy 1879 GTACCGGTTGACACACAGCCTTCTGTTGGATTTGAA 1919
Db 1261 GTCCGGTTGACACAAAATCTTCTGTTGGATTTGAA 1301
RESULT 14
US-10-820-714A-12
; Sequence 12, Application US/10820714A
; Publication No. US20050214922A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuoshi
; APPLICANT: Kobayashi, Toru
; APPLICANT: Sumitomo, Nobuyuki
; APPLICANT: Takimura, Yasushi
; APPLICANT: Sato, Tetsuo
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 251697US0
; CURRENT APPLICATION NUMBER: US/10/820,714A
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106709
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: *Bacillus* sp. KSM-Kp860
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: protease KP860
US-10-820-714A-12
Query Match 51.8%; Score 997; DB 9; Length 1302;
Best Local Similarity 85.4%; Pred. No. 5,1e-287;
Matches 1111; Conservative 0; Mismatches 190; Indels 0; Gaps 0;
Qy 619 AATGATGTTGCGCGTGAATTTGTCAAGCGATGTGCTCAGAGGAGCTACCGGTTGTAT 678
Db 1 AATGATGTTGCGCGAGTATTTGTGAAGCGATGTGCAACAGGAGCTACCGGTTGTAT 60
Qy 679 GGAAGAAGACAGATGTAGCGGTTGCCGATACAGGCTTGTATCAGGTGCAATGACAGT 738
Db 61 GGAAGAAGACAGATGTGTGCGAGTCCGATCTGATGATGATGATGATGATGAT 120
Qy 739 TCGATGCAATGAAAGCTTCCGCGGAAATTAATCTGATTAATGATGATGATGATGAT 798
Db 121 TCGATGCAATGAAAGCTTCCGCGGTAATAATGATGATGATGATGATGATGAT 180
Qy 799 AATGCAATGATACGAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 858
Db 181 AATGCAATGATACGAACGTGATGATGATGATGATGATGATGATGATGATGATGAT 240
Qy 859 TCCACTAATAAAGAAATGGCGCTCAGGGGAATCTAGTCTTCATCTATCTATGATAGC 918
Db 241 GCATGCAATTAAGAAATGGCACTCAGAGCAATCTGGTTCATCTATCTATGATAGC 300
Qy 919 GGTGGGGGACTTGAAGACTACCTTCGAATCTGCAACCTTATTCAGCAAGCATACAGT 978
Db 301 AGTGGGGGCTTGAAGGCTTGGCTTCATCTGCAACCTTATTCAGCAAGCATACAGT 360
Qy 979 GCTGTGCGCAATTCATCAAACTCTGGGGAGCAGCATGTAATGGGCTTACACACA 1038
Db 361 GCAGGTGCGCAATTCATCAAACTCTGGGGAGCAGCGGTGTAATGGGCTTACACACA 420
Qy 1039 GATTCCAGAAATGTGATGATCTATGTGCGCAAAATGATGACATCTTTCCTGCTGC 1098
Db 421 GATTCCAGAAATGTGATGATCTATGTGAGAAATGATGATGACATCTTTCCTGCTGC 480
Qy 1099 GGGAAATGAAGACCGAACCGGGAACCATCATGTCACAGGACAGCTTAAATGCAATA 1158

Db 481 GGGAAATGAAGGCGCGAAGCGGATGACATGACCTGATACGGCTTAAAAACCCATC 540
Qy 1159 ACAATCGAGAGCTACGGAACAACTCCGCCAGCTTGGGTCTTATGCGGACATATCAAC 1218
Db 541 ACTGTCGGCGCAACGGAACAACTGCGTCCAGCTTCGATTCCTATGACATATATTAAC 600
Qy 1219 CATGTGGCAAGTTCCTTCACTGATGACCGGACAAAGAGATGAGCGGATCAACCGGATGTC 1278
Db 601 CAGCTTGCACAGTCTCTCCGTCGCGCCGGAACAAAGATGAGGCGGAAATCAAGCTATGTC 660
Qy 1279 ATGGACCGGGAAAGCTTCACTATCAAGCAAGATCTTCTTGCACACCGGATTCCTCTTC 1338
Db 661 ATGGCGCAAGGAGCATACATTTTATCAGGAGATCTTCTTGCACACCGATTCCTCTTC 720
Qy 1339 TGGCGGAACCATGACAGTAATATGATATCATGGGTGGAACGTCATGCGTACACCGATC 1398
Db 721 TGGCGGAATCATGACAGCAAAATATGCTATATGGGTGGAACGTCATGCGCAACCGAAT 780
Qy 1399 GTTGTGGAACGTTGCGCAGCTTGTGAGCATTTTGTGAAAAAGAGGAGCATCACCA 1458
Db 781 GTTGTGGAATGTTGCAAGCTCCGTGAGCATTTTGTGAAAAATGAGGAATCACTCT 840
Qy 1459 AAGCTTCTCTATTTAAAGCGGACATGATGCGGTCGAGCTGACATCGGCTTGGCTAC 1518
Db 841 AAGCTTCTCTATTTAAAGCGCTTGAATGCGGTCGATGATGTTGGATGGGTTAT 900
Qy 1519 CCGAACGCTTACCAAGATGCGGAGCATGATGATGATTAATCCCTGATCGTTGCTTAT 1578
Db 901 CCGAACGGAACCAAGATGCGGCGAGTGAACCTGGAATTAATCTTGAACGTTGCTTAT 960
Qy 1579 GTGAACGAGTCAGTCTCTATCAACGCAAGCAAAAGAGGAGTACTCGTTTACTGCTAT 1638
Db 961 GTGAACGATCAAGTGCCTTATCAATGACCAAAAGAGGACATATACCTTTATGCAACG 1020
Qy 1639 GCGCGCAAGCTTTGAAAAATCTCCGTGATGCTGATGCCCCCTGCGAGCAACATGCT 1698
Db 1021 GCGCGCAAGCATTTGAAAATCTCCGTGATGCTGATGCCCCCTGCGAGCAACATGCT 1080
Qy 1699 TCCCTTAAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758
Db 1081 TCTGTAACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Qy 1759 GTGAAGAAATGACTTACTTGGCATATGATGATGATGATGATGATGATGATGATGATGAT 1818
Db 1141 GTCCGGAATGACTTCTGACACCATTTGATCAATGATGATGATGATGATGATGATGAT 1200
Qy 1819 AATGATTTATTTATGACCAACAAAGCGGAGCTATGATGATGATGATGATGATGATGAT 1878
Db 1201 AATGATTTATTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Qy 1879 GTACCGGTTGACCAAGACCTTCTGTTGGCAATTTGAA 1919
Db 1261 GTCCGCTTGAACCAAAACTTCTGTTGGCAATTTGAA 1301

RESULT 15

US-10-820-712A-21
; Sequence 21, Application US/10820712A
; Publication No. US20050026804A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuoshi
; APPLICANT: Izawa, Yoshifumi
; APPLICANT: Kodayashi, Toru
; APPLICANT: Koyama, Shingo
; APPLICANT: Sato, Tsuyoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 251701-USO
; CURRENT APPLICATION NUMBER: US/10/820,712A
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106708
; PRIOR FILING DATE: 2003-04-10

; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Bacillus NCIB12289
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: Gene name Al, protease Al
US-10-820-712A-21
Query Match 44.4%; Score 854.6; DB 8; Length 1302;
Best Local Similarity 76.6%; Pred. No. 2.2e-244;
Matches 1022; Conservative 0; Mismatches 279; Indels 0; Gaps 0;
Qy 619 AATGATGTTGCGCGTGAATTTGCAAGCGGATGCTCAGAGCAGCTACGGGTTGAT 678
Db 1 AACGATGTTGCAAGGCAATTTGAAAGCGATGTCAGAGCAGCTATGTTATAT 60
Qy 679 GGAACAGACAGATGCTAGCGGTTGCGGATCAGGCTTGTATCAGGTCGATACAGT 738
Db 61 GGGCAAGGCAAGTGTGCAATGCGGATCCGATCCGATGATATGCTTGGAAACAAAC 120
Qy 739 TCGATGATGAAGCGCTTCCGCGGAAATTAATGATATATGATGATGGAAGCAAT 798
Db 121 TCGATGATGAAGCGCTTCCGCGGAAATTAATGATATATGCTTGGAAACAAAC 180
Qy 799 AATGCAATGATACGAATGCTCATGATGCGATGCTGCTCGTATTAAGAAACGGC 858
Db 181 AAGCCATGATTCAAACCGGACAGGAGCGATGTTGCGGATCGTTTGAAGACGT 240
Qy 859 TCCACTATTAAGAAATGCGGCTCAGGCAATGATGCTTCCATCTATATGATGAC 918
Db 241 ACTTGAATTAAGAAATGCGCTCAGGCAATGATGCTTCCATCTATATGATGAC 300
Qy 919 GATGAGGAGCTTGAAGGATCTTCCGATGCTGAAACCTTATCAGGCAAGTACAGT 978
Db 301 AATGCGGATTTGCGGAGCTTCCGATGCTTCCGATGCTTCCGATGCTTCCGATGCT 360
Qy 979 GCTGTGCGAATTTCAATCAATCTCCGCGGAGCAAGTGAATGCGGCTTACAAACA 1038
Db 361 GCGGTCGCAATTCATACGATGCTTCCGATGCTTCCGATGCTTCCGATGCTTCCGAT 420
Qy 1039 GATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
Db 421 GATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Qy 1099 GGAATGAAGGACGAACCGGCGAACCATGATGACAGGCAAGCTAAATGCAATA 1158
Db 481 GGTAAAGGCGGAGGCGGAGGCGGACATGATGCTCTGATACCGGAGATGCTATC 540
Qy 1159 ACAATCGAGAGCTACGGAACAACTCCGCCAGCTTGGGTCTTATGCGGACATATCAAC 1218
Db 541 ACAATCGAGAGCAACGAACAACTCCGCCAGCTTGGATCTTATGCGGACATATCAAC 600
Qy 1219 CATGTGGCAAGTTCCTTCACTGATGACCGGACAAAGAGATGAGCGGATCAACCGGATGTC 1278
Db 601 CATGTGGCAAGTTCCTTCCGTCGCGCCGGAACAAAGATGAGCGGATCAACCGGATGTC 660
Qy 1279 ATGGACCGGGAAAGCTTCACTATCAAGCAAGATCTTCTTGCACACCGGATTCCTCTTC 1338
Db 661 ATGGCGCAAGGAGCATACATTTTATCAGGAGATCTTCTTGCACACCGATTCCTCTTC 720
Qy 1339 TGGCGGAACCATGACAGTAATATGATATCATGGGTGGAACGTCATGCGTACACCGATC 1398
Db 721 TGGCGGAATCATGACAGCAAAATATGCTATATGGGTGGAACGTCATGCGCAACCGAAT 780
Qy 1399 GTTGTGGAACGTTGCGCAGCTTGTGAGCATTTTGTGAAAAAGAGGAGCATCACCA 1458
Db 781 GTTGTGGAATGTTGCAAGCTCCGTGAGCATTTTGTGAAAAATGAGGAATCACTCT 840
Qy 1459 AAGCTTCTCTATTTAAAGCGGACATGATGCGGTCGAGCTGACATCGGCTTGGCTAC 1518

Db 841 AAACCATCTTGTGTAAGCAGCTCTATATGCCGAGCACTGATATCGGTCTTGCTAT 900
QY 1519 CCGAAGGTAAACGAAGATGGGACGATGACATTTGATAATCCCTGAACGTTGCTAT 1578
Db 901 CCGAGTGGAAACGAAGATGGGAGATGACATTTGACAACTTAAATGTAGCTTTC 960
QY 1579 GTGAACGAGTCCAGTCTCTATCCACGAGCAAAAAGCGATCTGTTACTGCTACT 1638
Db 961 GTAAATGAAACGAGCTCTTATCTATACCAAAAGCTAGTATTCATTTACTGCAAA 1020
QY 1639 GCGGCAAGCTTTGAAAATCTCCGTGTATGCTGTATGCCCTGCGAGCACACTGCT 1698
Db 1021 TCAGGCAAACTTTGAAGATTTCAATTGGTTGGTCTGATGCACGCGAAGTACTTCGCA 1080
QY 1699 TCGGTAAAGCTTGTCAATGATCTGAGACTTGTCAATACGCTCCAAATGSCACAAGTAT 1758
Db 1081 TCGGTACATTTGGTGAATGATCTGATCTGTGATTAAGCTCCAAATGGAACAAAGTAT 1140
QY 1759 GTAGAAATGACTTTACTTCCGATTAATGATTAATGAGATGAGCGCAATTAAGTAGAA 1818
Db 1141 GTTGAACGACTTTACTGCTCCCTATGATTAATTAATGAGATGAGCGCAATTAAGTAGAG 1200
QY 1819 AATGATTTTATTAATGACCAACAAGCGGACGATTAATTAATGAGTACAGGCTTAAAC 1878
Db 1201 AAGGTGTTTATCAATGCTCGCAAAAGGGAAGTATTAAGTGAAGTACAGGCTTAAAC 1260
QY 1879 GTACCGGTGAGCAACAAGCTTCTCGTTGGCAATTTGTAA 1919
Db 1261 GTTCCACAAGGCGCGAGCGTTCCTTGGCTATTGTGAA 1301

Search completed: April 8, 2006, 07:34:29
Job time : 1384.58 secs

This Page Blank (uspto)


```
RESULT 5
US-09-925-065A-562287/c
; Sequence 562287, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 562287
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-562287

Query Match
Best Local Similarity 2.1%; Score 39.6; DB 6; Length 417;
Matches 57; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 243 CCCGGCAATTAATACTCCATATTATCCATTCAATGACCAATTTTGAAGAAACAA 302
DB 230 CCCACATATTAATAATACATATTAATAATTAATAATTAATAATTAATAACAA 171
QY 303 ACAGCAGCTGGAATAAACAAGGGGCAA 328
DB 170 ACAGCAGATGTAAATGACAGAGCAA 145

RESULT 6
US-09-925-065A-124623/c
; Sequence 124623, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124623
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-124623

Query Match
2.1%; Score 39.6; DB 6; Length 514;
```

```
Best Local Similarity 66.3%; Pred. No. 0.76;
Matches 57; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 243 CCCGGCAATTAATACTCCATATTATCCATTCAATGACCAATTTTGAAGAAACAA 302
DB 422 CCCACATATTAATAATACATATTAATAATTAATAATTAATAATTAATAACAA 363
QY 303 ACAGCAGCTGGAATAAACAAGGGGCAA 328
DB 362 ACAGCAGATGTAAATGACAGAGCAA 337

RESULT 7
US-10-301-480-75030/c
; Sequence 75030, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75030
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-75030

Query Match
Best Local Similarity 2.1%; Score 39.6; DB 9; Length 527;
Matches 57; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 243 CCCGGCAATTAATACTCCATATTATCCATTCAATGACCAATTTTGAAGAAACAA 302
DB 434 CCCACATATTAATAATACATATTAATAATTAATAATTAATAATTAATAATTAACAA 375
QY 303 ACAGCAGCTGGAATAAACAAGGGGCAA 328
DB 374 ACAGCAGATGTAAATGACAGAGCAA 349

RESULT 8
US-10-301-480-688439/c
; Sequence 688439, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 688439
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-688439

Query Match
2.1%; Score 39.6; DB 10; Length 527;
Best Local Similarity 66.3%; Pred. No. 0.77;
```


NUMBER OF SEQ IN NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 841440
LENGTH: 656
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-841440

Query Match 1.9%; Score 37.4; DB 6; Length 656;
Best Local Similarity 57.1%; Pred. No. 3.8;
Matches 68; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 436 CCGATATACAGATTAATCCCGAGCTTTTCACAAAGGGGCATCAGAGCTTGTAAAGCA 495
DB 132 CCGAAGACATTAATCTTGTCAAACTTAATTAATCAAGACAGAGTTAAATGGA 73
QY 496 GTGGCGCTTGATACAAAGCAAAATTAAGAGTGCAATTAAGAGCATCGAACAAT 554
DB 72 GTGCAGCTTGACAAATTAATAAGAAAAAGAAAAAGTAAATCATGCATCAACTAAT 14

RESULT 13
US-09-925-065A-31269/c
Sequence 31269, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31269
LENGTH: 1766
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-31269

Query Match 1.9%; Score 37.4; DB 6; Length 1766;
Best Local Similarity 57.1%; Pred. No. 6.3;
Matches 68; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 436 CCGATATACAGATTAATCCCGAGCTTTTCACAAAGGGGCATCAGAGCTTGTAAAGCA 495
DB 178 CCGAAGACATTAATCTTGTCAAACTTAATTAATCAAGACAGAGTTAAATGGA 119
QY 496 GTGGCGCTTGATACAAAGCAAAATTAAGAGTGCAATTAAGAGCATCGAACAAT 554
DB 118 GTGCAGCTTGACAAATTAATAAGAAAAAGAAAAAGTAAATCATGCATCAACTAAT 60

RESULT 14
US-10-301-480-132507/c
Sequence 132507, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480

CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 132507
LENGTH: 1766
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-132507

Query Match 1.9%; Score 37.4; DB 9; Length 1766;
Best Local Similarity 57.1%; Pred. No. 6.3;
Matches 68; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 436 CCGATATACAGATTAATCCCGAGCTTTTCACAAAGGGGCATCAGAGCTTGTAAAGCA 495
DB 178 CCGAAGACATTAATCTTGTCAAACTTAATTAATCAAGACAGAGTTAAATGGA 119
QY 496 GTGGCGCTTGATACAAAGCAAAATTAAGAGTGCAATTAAGAGCATCGAACAAT 554
DB 118 GTGCAGCTTGACAAATTAATAAGAAAAAGAAAAAGTAAATCATGCATCAACTAAT 60

RESULT 15
US-10-301-480-745916/c
Sequence 745916, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 745916
LENGTH: 1766
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-745916

Query Match 1.9%; Score 37.4; DB 10; Length 1766;
Best Local Similarity 57.1%; Pred. No. 6.3;
Matches 68; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 436 CCGATATACAGATTAATCCCGAGCTTTTCACAAAGGGGCATCAGAGCTTGTAAAGCA 495
DB 178 CCGAAGACATTAATCTTGTCAAACTTAATTAATCAAGACAGAGTTAAATGGA 119
QY 496 GTGGCGCTTGATACAAAGCAAAATTAAGAGTGCAATTAAGAGCATCGAACAAT 554
DB 118 GTGCAGCTTGACAAATTAATAAGAAAAAGAAAAAGTAAATCATGCATCAACTAAT 60

Search completed: April 8, 2006, 03:55:08
Job time : 898.7 secs

This Page Blank (uspto)